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

METODY TEMPORÁLNÍHO TONE-MAPPINGU HDR VIDEA

TEMPORAL TONE-MAPPING METHODS FOR HDR VIDEO

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE — Bc. JIŘÍ STANĚK
AUTHOR

VEDOUCÍ PRÁCE — Ing. SVETOZÁR NOSKO
SUPERVISOR

BRNO 2019

Ústav počítačové grafiky a multimédií (UPGM) Akademický rok 2018/2019

Zadání diplomové práce

21624

Student: **Staněk Jiří, Bc.**
Program: Informační technologie Obor: Počítačová grafika a multimédia
Název: **Metody temporálního tone-mappingu HDR videa**
Temporal Tone-Mapping Methods for HDR Video
Kategorie: Zpracování obrazu
Zadání:

1. Prostudujte dostupnou literaturu týkající se HDR a prostudujte metody temporálního tone-mappingu HDR videa. Po konzultaci s vedoucím vyberte vhodnou metodu pro temporální tone-mapping HDR videa.
2. Proveďte návrh programu demonstrujícího použití metody vybrané v předchozím bodě zadání.
3. Implementujte demonstrační aplikaci, která na HDR video aplikuje vybranou metodu temporálního tone-mappingu pro zobrazení.
4. Diskutujte dosažené výsledky a navrhněte možné pokračování práce.

Literatura:

- Reinhard, Erik, et al. *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.

Při obhajobě semestrální části projektu je požadováno:

- Body 1 a 2.

Podrobné závazné pokyny pro vypracování práce viz http://www.fit.vutbr.cz/info/szz/

Vedoucí práce: **Nosko Svetozár, Ing.**
Vedoucí ústavu: Černocký Jan, doc. Dr. Ing.
Datum zadání: 1. listopadu 2018
Datum odevzdání: 22. května 2019
Datum schválení: 1. listopadu 2018

Abstrakt

Tato diplomová práce poskytuje teoretické náležitosti nutné pro pochopení komplexní problematiky vysokého dynamického rozsahu a řeší implementaci inovativní metody pro temporální mapování tónů videa. Nejprve je vysvětleno pořizování, ukládání a zpracování obrazu s vysokým dynamickým rozsahem. Následuje obecné uvedení mapování tonality, popis nežádoucích obrazových artefaktů a přehled nejvyspělejších existujících algoritmů. Další část se věnuje implementaci zvolené metody řešení a navrhuje několik dodatečných vylepšení původně vybraného algoritmu. Na základě uvedených informací vznikla aplikace s intuitivním grafickým rozhraním, kterou lze využít pro temporální mapování tónů HDR videa. Na závěr jsou dosažené výsledky demonstrovány a vizuálně porovnány s již existujícími metodami pro temporální mapování tónů videa.

Abstract

The diploma thesis provides the theoretical background needed to understand the complex topic of high dynamic range, and deals with the implementation of an innovative method for temporal video tone-mapping. At first, high dynamic range image acquisition, storage and processing are described. Afterwards, a general introduction to tone-mapping, description of undesirable image artifacts and an overview of existing state-of-the-art algorithms are presented. The part which follows focuses on the implementation of the chosen method and proposes several additional enhancements of the initially selected algorithm. Based on the information provided, an application with an intuitive graphical interface, which can be used for temporal HDR video tone-mapping, has been created. Finally, the the achieved results are demonstrated and visually compared with existing methods for temporal video tone-mapping.

Klíčová slova

mapování tónů, vysoký dynamický rozsah, HDR, digitální video, temporální koherence, temporální adaptace

Keywords

tone-mapping, hygh dynamic range, HDR, digital video, temporal coherency, temporal adaptation

Citace

STANĚK, Jiří. *Metody temporálního tone-mappingu HDR videa.* Brno, 2019. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Svetozár Nosko

Metody temporálního tone-mappingu HDR videa

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Ing. Svetozára Noska. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

. . . . . .
Jiří Staněk
20. května 2019

Poděkování

Rád bych poděkoval panu Ing. Svetozáru Noskovi za odbornou pomoc a čas, který mi v průběhu psaní práce věnoval.

Obsah

1 Úvod **3**

2 Teoretický úvod a důležité pojmy **4**
2.1 Základní fyzikální principy 4
2.1.1 Světlo 4
2.1.2 Vnímání barev člověkem 5
2.1.3 Fotometrické veličiny 6
2.2 Digitální fotografie a video 6
2.2.1 Komponenty digitálních fotoaparátů a videokamer 6
2.2.2 Expozice 8
2.2.3 Dynamický rozsah 9

3 Obraz s vysokým dynamickým rozsahem **11**
3.1 Formáty pro HDR snímky 11
3.1.1 Radiance HDR 11
3.1.2 OpenEXR 12
3.2 Formáty pro HDR video 12
3.2.1 Standardy pro HDR video 13
3.2.2 Formáty kódování HDR videa 13
3.3 Pořizování HDR obrazu 15
3.3.1 Skládání HDR obrazu pomocí fúze expozic 16
3.3.2 Pořizování vstupních snímků pro fúzní algoritmy 17

4 Mapování tónů **20**
4.1 Extrakce základové a detaily nesoucí vrstvy 21
4.1.1 Gaussův filtr 21
4.1.2 Bilaterální filtr 22
4.1.3 Řízený filtr 22
4.2 Mapování rozsahu podle tónovací křivky 24
4.2.1 Lokální adaptace tónovací křivky 24
4.3 Rozdíly mezi mapováním tónů statických snímků a videa 24
4.3.1 Obrazové artefakty 25
4.3.2 Temporální koherence videa 27
4.3.3 Viditelnost a zesílení šumu 28
4.4 Přehled vybraných operátorů pro mapování tónů videa 28
4.5 Operátor ELTM 30
4.5.1 Obecný popis algoritmu ELTM 30
4.5.2 Stanovení relativní luminance 32

4.5.3 Dekompozice obrazu . . . 32
4.5.4 Komprese rozsahu kontrastu základové vrstvy . . . 33
4.5.5 Zpracování detailů . . . 34
4.5.6 Komprese tónů . . . 35
4.5.7 Obnovení detailů a barev . . . 36
4.5.8 Temporální zpracování snímků . . . 36

5 Návrh a implementace aplikace 40
5.1 Požadavky na funkcionalitu aplikace . . . 40
5.2 Požadavky na uživatelské rozhraní . . . 41
5.3 Nástroje zvolené pro implementaci . . . 41
5.4 Implementace funkcionality . . . 42
5.4.1 Temporální předzpracování snímků . . . 42
5.4.2 Extrakce relativní luminance . . . 45
5.4.3 Dekompozice obrazu a zpracování detailů . . . 46
5.4.4 Komprese rozsahu základové vrstvy . . . 47
5.4.5 Obnovení barev . . . 47
5.4.6 Temporální zpracování snímků a zápis videa . . . 47
5.5 Grafické uživatelské rozhraní . . . 50
5.5.1 Implementace grafického uživatelského rozhraní . . . 51

6 Ukázka a zhodnocení výsledků 53
6.1 Dopad statických parametrů na výsledný obraz . . . 53
6.2 Ukázka vlivu rychlosti temporální adaptace . . . 55
6.3 Vliv referenčního snímku na výsledné video . . . 56
6.4 Ukázka výsledků a srovnání s nejvyspělejšími operátory . . . 58
6.5 Rychlost implementované aplikace . . . 62

7 Závěr 63

Literatura 64

Přílohy 71

A Obsah přiloženého paměťového média 72

Kapitola 1

Úvod

Existuje mnoho odvětví informačních technologií, jež zahrnují práci s digitálním obrazem. Ten může být zachycen fotoaparátem, kamerou nebo kupříkladu vykreslen počítačovým programem. V posledních letech byl kladen velký důraz na kontinuální zvyšování rozlišení pořizovaného obrazu, ovšem nyní se pozornost stále více upíná na možnosti samotné reprezentace obrazových dat. V dnešní době stále dominuje přístup, kde je pro každou z barevných složek, typicky pro červenou, zelenou a modrou, vyhrazen pouze jeden byte. Každý obrazový bod tedy může nabývat jeden ze 16,7 milionů možných odstínů barev. Takové číslo se sice může na první pohled jevit jako dostatečné, nicméně je nutné si uvědomit, že jeden byte umožňuje každé barvě nabývat pouze 256 různých hodnot, což je oproti lidskému vidění značně omezený rozsah [65].

HDR je obor zpracování obrazu, jenž se zaměřuje na pořizování, zpracování a zobrazování obrazu s vysokým dynamickým rozsahem. Jeho primárním cílem je dosažení co nejvěrnější reprodukce fyzické hodnoty světla na číslicových zařízeních [55]. Vzhledem ke stále rostoucímu výkonu výpočetních zařízení a osobních počítačů se zvyšují také nároky uživatelů na kvalitu a věrnost reprodukovaného obrazu. HDR tematika se tak stává velmi aktuálním předmětem zkoumání a vývoje v akademické i komerční sféře. Většina domácností prozatím nedisponuje zařízením, jež by dokázalo nativně zobrazovat HDR obraz. Kvůli této skutečnosti vzniká řada algoritmů pro kompresi obrazu s vysokým dynamickým rozsahem do rozsahu, který je možné zobrazit na klasických obrazovkách s běžným dynamickým rozsahem. Taková komprese se souhrnně nazývá mapování tónů. Jedná se o poměrně komplexní úlohu, zvláště v případě videa, kde je nutné uvážit celou řadu problémů spojených s temporální dimenzí obrazu.

Prvním cílem této práce je vysvětlit úskalí spojená s mapováním tónů videa. Oproti statickým snímkům je nutné zohlednit mnoho faktorů, jež mohou negativně ovlivnit vzájemnou koherenci jednotlivých snímků výsledného videa. Druhým cílem je implementovat algoritmus, který dokáže adaptivně modifikovat své výpočty v časové doméně, čímž zajistí temporální stabilitu výstupního videa. Kapitola 2 přibližuje základní fyzikální principy spojené s vnímáním obrazové reality a popisuje pojmy z oblastí digitální fotografie a zpracování obrazu. Následuje kapitola 3, která se věnuje ukládání a pořizování snímků s vysokým dynamickým rozsahem. Obecný princip mapování tónů a popis nejzajímavějších existujících algoritmů používaných pro video uvádí kapitola 4. Návrh a implementaci výsledného programu přibližuje kapitola 5, která se věnuje i grafickému uživatelskému rozhraní aplikace. V předposlední kapitole jsou demonstrovány dosažené výsledky a výstupní snímky jsou vizuálně porovnány s existujícími metodami pro mapování tónů videa. Závěrečná část textu shrnuje celou práci a představuje směr, kterým by se mohla dále v budoucnosti ubírat.

Kapitola 2

Teoretický úvod a důležité pojmy

Tato kapitola vysvětluje nejdůležitější pojmy potřebné pro hlubší pochopení principů používaných při práci s obrazem o vysokém dynamickém rozsahu. Jedná se zejména o přiblížení základních fyzikálních principů souvisejících s přenosem světla a také o uvedení fotografického názvosloví, které je hojně používáno v dalších kapitolách tohoto textu.

2.1 Základní fyzikální principy

Hlavní motivací pro používání obrazu s vysokým dynamickým rozsahem je co nejpřesnější zachycení světla a jeho následná reprodukce z pohledu pozorovatele. Lidská zraková soustava umožňuje spojitě rekonstruovat obraz prostředí na základě světla odraženého od objektů v prostoru [61].

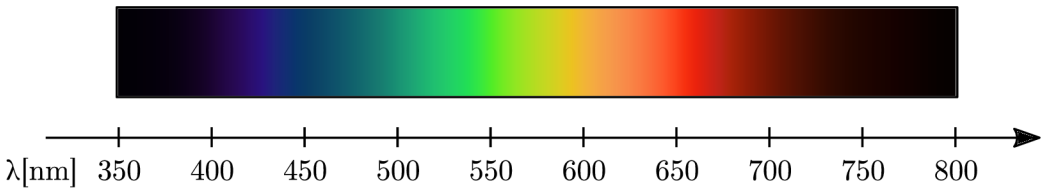
2.1.1 Světlo

Z pohledu vlnové optiky označujeme světlo jako člověkem viditelnou část spektra elektromagnetického vlnění. Jedná se o děj, jenž ke svému šíření nepotřebuje žádné látkové prostředí a šíří se tedy například i vakuem [63]. Světlo, obdobně jako další druhy vlnění, charakterizuje jeho vlnová délka λ, která vyjadřuje vzdálenost mezi nejbližšími body postupného periodického vlnění ve stejné fázi.

Další charakteristickou veličinou je frekvence f, která udává počet jednotlivých period vlny za jednotku času. Pomocí rychlosti světla v určitém prostředí v pak lze vypočítat vlnovou délku světla pomocí vztahu:

$$\lambda = \frac{v}{f}. \tag{2.1}$$

Lidský zrak dokáže vnímat světlo o přibližné vlnové délce 390 nm až 790 nm [63]. Tento rozsah je pouze orientační, protože je ovlivněn individualitou každého člověka. Přibližné schéma vnímání barev podle vlnové délky světla je ilustrováno obrázkem 2.1.

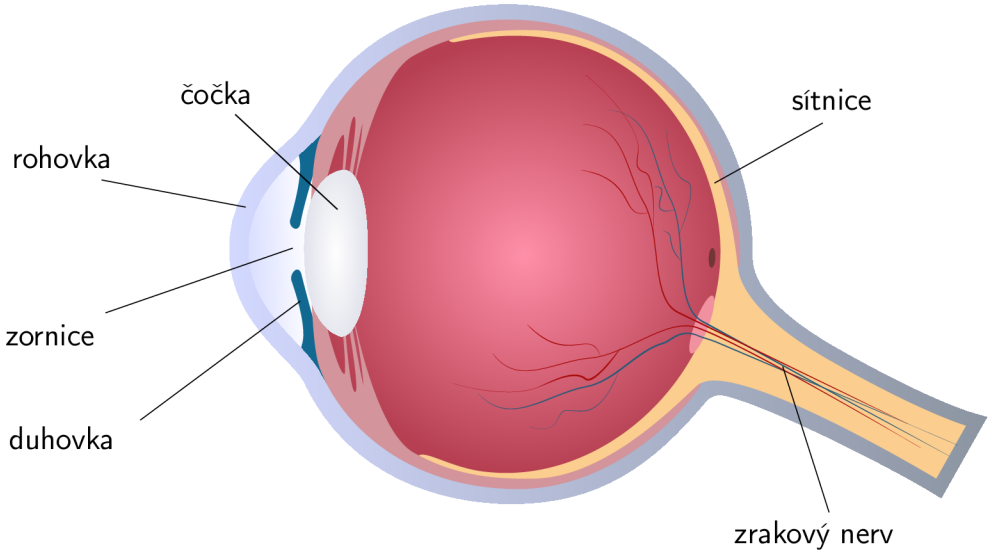


Obrázek 2.1: Spektrum viditelné lidským okem [63].

2.1.2 Vnímání barev člověkem

Zrakový vjem, tedy proces transformace elektromagnetického světelného vlnění na elektrické signály proudící do mozku, je velmi složitý proces. Světlo proniká do oka skrz rohovku a jeho množství je regulováno duhovkou, která dokáže smrštit či roztáhnout svoji středovou štěrbinu podle aktuálního osvětlení prostředí. Tato štěrbina se nazývá zornice. Propuštěné paprsky dále prochází čočkou, která umožňuje člověku ostřit na potřebnou vzdálenost. Fyzikálním principem tohoto procesu zvaného *akomodace* je změna ohniskové vzdálenosti[1] čočky. Soustava malých svalů upnutých na čočku umožňuje změnu jejího zakřivení a optické mohutnosti, čímž je efektivně proveden posun ohniska[2] do místa, kde se nachází sítnice. S přibývajícím věkem jedince se postupně zhoršuje pružnost čočky, což se projevuje zhoršenou schopností ostřit na blízké předměty.

Samotnou transformaci světelného vlnění na elektrický signál zprostředkovává sítnice, což je tenká vrstva na vnitřní straně oka obsahující speciální fotocitlivé buňky. Prvním druhem těchto citlivých buněk jsou tzv. tyčinky, kterých je na sítnici přibližně 120 milionů a které umožňují vidění při nízké intenzitě světla. Neumožňují však rozpoznávat barvy. Buňky druhého typu se nazývají čípky a umožňují barevné vidění v osvětlených prostorech. Na sítnici se nachází přibližně 20krát méně čípků než tyčinek, tedy okolo 6 milionů [34]. Následuje obrázek 2.2 popisující fyzickou stavbu lidského oka.



Obrázek 2.2: Elementární anatomie lidského oka [34].

Zjednodušeně řečeno, fotocitlivé buňky reagují na dopad světla chemickou reakcí, která vytváří elektrické impulsy. Ty pokračují skrze zrakový nerv dále do týlního laloku mozku, kde dochází k vytvoření samotného zrakového vjemu. Zajímavostí je, že oko je nejcitlivější na světlo o vlnové délce přibližně 555 nm, což odpovídá zelené barvě [84]. Zelená barva se nachází přibližně ve středu spektra viditelného světla a při dopadu na sítnici je zapojeno největší množství čípků.

[1] Ohnisková vzdálenost – vzdálenost mezi zakřivenou čočkou a jeho ohniskem.
[2] Ohnisko – bod, kde se světelné paprsky procházející čočkou setkávají.

2.1.3 Fotometrické veličiny

Fotometrie je fyzikální věda zkoumající světlo z hlediska jeho efektu na lidskou zrakovou soustavu. Jedná se o disciplínu, kterou lze vyložit jako paralelní vědu vůči radiometrii[3], protože obě využívají obdobné principy. Oproti radiometrii se fotometrie zabývá pouze okem viditelnou částí elektromagnetického záření. Rozdíly jsou patrné i v názvosloví, a proto je vhodné uvést nejzásadnější fotometrické veličiny z pohledu této práce:

- **Světelný tok**: viditelná světelná energie vyzářená zdrojem za jednotku času.
  - symbol veličiny: Φ
  - jednotka: lm (lumen)
- **Svítivost**: světelný tok vyzařován bodovým zdrojem pod určitým úhlem do prostou.
  - symbol veličiny: I
  - jednotka: cd (kandela)
- **Luminance (také jas)**: měrná veličina svítivosti vztažená na danou plochu.
  - symbol veličiny: L
  - jednotka: cd/m^2
- **Osvětlení**: světelný tok dopadající na plochu o určité velikosti.
  - symbol veličiny: E
  - jednotka: lx (lux)

2.2 Digitální fotografie a video

Aby bylo možné řádně vysvětlit problematiku dynamického rozsahu v kontextu zpracování obrazu, je nutné nejprve uvést základní pojmy z oblasti digitální fotografie. Jedná se o popis dílčích komponent fotoaparátu a nastavitelných parametrů, se kterými lze manipulovat při pořizování snímků či videa. V závěru této podkapitoly jsou vysvětleny dva velice důležité pojmy – expozice a dynamický rozsah.

2.2.1 Komponenty digitálních fotoaparátů a videokamer

Digitální zařízení pro pořizování obrazu jsou do jisté míry inspirována lidským viděním. Už od dob klasických „analogových“ fotoaparátů se spočívá princip jejich fungování v osvětlování plochy citlivé na světlo, která transformuje světelné paprsky na obrazovou informaci. Před samotným dopadem světla na tuto plochu je nutné nasměrovat a propustit pouze část paprsků podle požadavků na výsledný obraz. K tomu slouží objektiv aparátu a závěrka, která se nachází na těle fotoaparátu před snímacím čipem.

Pro klasické aparáty se jako citlivá vrstva používá jednorázový fotografický film, na kterém po vystavení světlu vzniká latentní obraz, který je nutné chemicky upravit tak, aby z něj vznikl obraz okem viditelný. Tento proces se nazývá *vyvolání* filmu. Naproti tomu digitální zařízení jako plochu citlivou na světlo obsahují speciální senzory, které zajišťují převod obrazu na číslicovou reprezentaci.

[3] Radiometrie – věda zkoumající elektromagnetické záření.

Objektiv

Světlo odražené scénou nejprve vstupuje do objektivu fotoaparátu, což je komponenta, jejíž účelem je usměrnit paprsky vhodným způsobem na senzor aparátu. Mobilní a kompaktní zařízení mívají typicky objektiv s jednou čočkou, který bývá jejich pevnou součástí. Dražší přístroje často umožňují výměnu objektivu. Aby mohl objektiv podporovat optické přiblížení, tzv. *optický zoom*, musí obsahovat soustavu několika čoček, díky kterým je umožněn posun jeho ohniskové vzdálenosti vůči snímači. Přiblížení či oddálení snímané scény je pak možné realizovat bez ztráty kvality výsledného snímku.

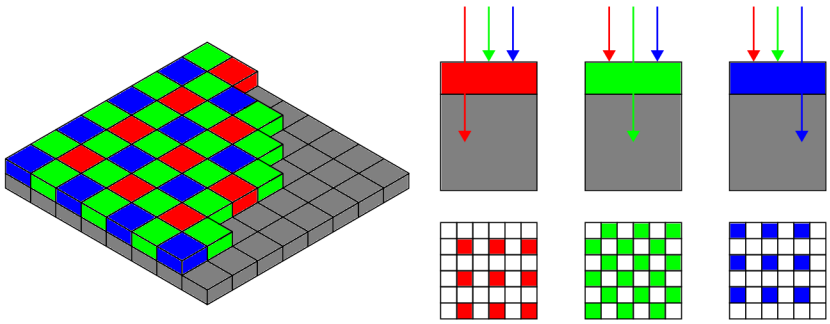
Důležitou součástí objektivu je také clona, která pomocí smrštění nebo roztažení dokáže regulovat množství světla propuštěného do těla fotoaparátu. Funguje tedy principiálně jako lidská duhovka. Clona má také vliv na hloubku ostrosti obrazu, což je rozsah vzdáleností ve scéně, který se na výsledném snímku jeví lidskému oku jako ostrý.

Závěrka

Před samotným senzorem se v těle fotoaparátu nachází zařízení zvané závěrka. Jde ve své podstatě o stínítko, které svým pohybem po stisknutí spouště umožní osvětlení snímacího čipu. Nastavením rychlosti závěrky lze ovlivnit dobu, po kterou je senzor vystaven světlu. V dnešní době se používá zejména štěrbinový typ závěrky, kde expozici snímače zajišťují dvě lamely, které se pohybují vertikálně. Na počátku expozice se první lamela otevře a po požadované expoziční době druhá lamela zakryje snímač. Při krátkém expozičním čase se druhá lamela uzavírá velmi rychle, a světlo je tak propouštěno pouze velmi úzkou horizontální štěrbinou, jež postupně osvětluje senzor směrem shora dolů. To vede k obrazovému artefaktu, kdy dochází k deformaci objektů, jež se ve scéně velmi rychle pohybují.

Snímací čip

Základní částí snímacího čipu je plocha vystavovaná světlu, která je tvořena maticí buněk s citlivými fotodiodami. Fotografická literatura označuje takové buňky jako tzv. *sensely* (z angl. *sensitive cell*) [37]. Sensel je schopen snímat světelnou intenzitu a následně ji převádět na elektrický signál. Na povrchu každého z nich je barevný filtr propouštějící pouze určitou vlnovou délku světla. Typicky jsou filtry uspořádány do mřížky tak, že dohromady tvoří tzv. *Bayerovu masku*. Filtry propouštějící zelenou barvu jsou zastoupeny v dvojnásobném počtu oproti modrým a červeným, díky čemuž je reflektována fyziologie lidského vidění (viz 2.1.2). Obecné schéma Bayerovy masky je uvedeno na následujícím obrázku 2.3:



Obrázek 2.3: Bayerova maska. Schémata vytvořil Colin M.L. Burnett pro Wikipedii[4].

[4] https://wikipedia.org/

Pro každý sensel je nasnímána intenzita osvětlení, která je dále převedena na elektrický signál. Následně proběhne pro každou barevnou složku interpolace bodu s jeho okolím, čímž vznikne jeden obrazový bod výsledného snímku. Typicky je do interpolace zahrnuto kromě samotného centrálního bodu jeho osm nejbližších sousedů. Tento proces bývá označován jako *debayerizace* a existuje několik *debayerizačních* metod, které mohou interpolovat barvy z různě velikého okolí, nebo například mohou používat více průchodů obrazovou maticí [49].

Dvě nejrozšířenější technologie senzorů, používané v moderních fotoaparátech a kamerách, jsou *Charge Coupled Device*, zkráceně *CCD*, a *Complementary Metal Oxide Semiconductor* neboli *CMOS* [45]:

- **CCD** snímač pracuje tak, že během vystavení buňky světlu v ní nashromáždí určité množství elektronů odpovídající intenzitě osvětlení. Náboje jsou následně z buněk odvedeny do CCD obvodů, kde se dále musí zpracovat. Postupně náboje prochází horizontálně napříč čipem skrze speciální posuvné registry. Takový posun nastane vždy synchronně vůči hodinovému signálu. V každém taktu jsou z okrajového sloupce odeslány náboje z jednotlivých registrů do výstupního zesilovače a následně proběhne převod na číslicovou podobu pomocí analogově digitálního převodníku [52].

- **CMOS** čip oproti CCD zařízením funguje tak, že každá buňka disponuje vlastními tranzistory pro transformaci náboje na digitální reprezentaci. Vzhledem k tomu, že výroba klasických mikročipů a CMOS senzorů je ve svém principu velmi podobná, stal se z technologie CMOS během posledních dvou dekád dominantní přístup na poli obrazových snímačů. Díky kontinuální miniaturizaci a zvyšování kvality produkovaných senselů se dnes příznivě vyvíjí nejen pořizovací cena, ale i reprodukční vlastnosti a spotřeba energie čipů CMOS. Typickým nedostatkem oproti snímačům typu CCD je vyšší výskyt šumu na výsledných fotografiích. Tento problém se ale postupně daří řešit díky rostoucí výrobní kvalitě, zvyšování rozlišení senzorů a také přidáváním speciálních obvodů pro explicitní redukci šumu [51].

2.2.2 Expozice

Pro charakterizaci vystavení snímače světlu se používá pojem *expozice*. Ve své podstatě jde o koordinaci jednotlivých součástí fotografické optiky tak, aby propustily na snímač světlo požadovaným způsobem. Na konečnou expozici mají vliv tři zásadní parametry:

- **Expoziční čas**: čas, během kterého závěrka umožní průchod světla na senzor. Typicky je udáván v sekundách a jejich zlomcích. Čím vyšší je rychlost závěrky, tím kratší je expoziční čas, kdy je senzor je vystaven světlu.

- **Velikost otvoru clony**: velikost štěrbiny ovlivňuje množství světla, jež dopadá na závěrku, která jej propouští dále na snímač. Ovlivňuje nejen expozici, ale také výslednou hloubku ostrosti snímku.

- **Citlivost ISO**: mění citlivost samotného snímače, přičemž v digitálních fotoaparátech se jedná o zesílení zachyceného signálu. Typicky lze na aparátu nastavit ISO citlivost na hodnoty 50, 100, 200, 400, 800, 1200, 1600, atd. Dnešní zařízení umožňují snímat s citlivostí až několika desítek tisíc. V praxi se ovšem citlivost ISO vyšší než 1600 používá velmi zřídka, protože s rostoucím zesilováním signálu roste ve velké míře výskyt šumu na produkovaných snímcích.

Kombinací uvedených parametrů lze nastavit požadovanou expozici. Při zvětšení průměru clonové štěrbiny je objektivem propuštěno více světla a je nutné snížit rychlost závěrky, a tedy zvýšit expoziční čas pro získání stejně exponovaných snímků. Zachycený signál je následně zesílen podle zvolené citlivosti, což může byt vhodné například při fotografování či natáčení nedostatečně osvětlené scény. Matematické vyjádření hodnoty expozice udává jednotka EV (z angl. *exposure value*), kterou lze vypočítat podle vzorce 2.2:

$$EV = log_2 \frac{N^2}{t}, \tag{2.2}$$

kde N je clonové číslo, které bývá často zapisováno ve tvaru f/N, a t je expoziční čas [68]. Hodnota f, používaná pro reprezentaci clonového čísla, vyjadřuje ohniskovou vzdálenost objektivu. Je patrné, že různé nastavení parametrů můžeme vést ke stejným expozičním hodnotám. Např. pro clonové číslo $f/5.6$ a expoziční čas $1/4$ s je hodnota EV rovna 7, stejně tak jako pro parametr clony $f/4$ a čas vystavení o délce $1/8$ s.

Je-li do výpočtu zahrnuta i citlivost ISO, lze hovořit o veličině zvané LV (z angl. *light value*) [33], kterou lze vypočítat podle rovnice 2.3:

$$LV = EV + log_2 \frac{ISO}{100}. \tag{2.3}$$

2.2.3 Dynamický rozsah

V obecné rovině lze dynamický rozsah vysvětlit jako rozdíl mezi nejmenším a největším bodem oboru hodnot libovolné veličiny. V oblasti zpracování obrazu jde o poměr mezi největším rozdílem jasů a lze hovořit o dynamickém rozsahu v kontextu pořizování, zpracování i reprodukce obrazu. Pro kvantifikaci se používá jednotka EV popsaná výše. Dynamický rozsah je tedy rozdíl mezi nejtmavším a nejjasnějším bodem, které se ve stejnou chvíli vyskytují na různých místech pozorované scény, resp. fotografie. Nezáleží na celkové světelnosti scény, ale na rozpětí světel a stínů.

Obrázek 2.4: Dynamický rozsah jednotlivých oblastí se může výrazně lišit od globálního rozsahu celé scény. Fotografii pořídil Spencer Gurley[5].

[5] https://pexels.com/@spencergurley

Jednotlivé vyznačené části scény na obrázku 2.4 mají samy o sobě malý dynamický rozsah, protože rozpětí mezi nejsvětlejším a nejtmavším bodem v rámci jednotlivých oblastí není velký. Celá scéna však disponuje velkým rozdílem mezi nejjasnějším a nejtmavším bodem, díky čemuž má o poznání vyšší dynamický rozsah než její dílčí části.

Běžné digitální aparáty typicky umožňují zachytit rozsah 6 až 14 EV [12]. Lidské oko je díky aktivní adaptaci schopno zobrazovat škálu přibližně 24 EV, a proto se fotoaparátem pořízené snímky označují jako snímky s nízkým dynamickým rozsahem. Digitální optika se nedokáže vypořádat s nepříznivými světelnými podmínkami, a dochází tak ke ztrátě detailů v oblastech, které se svoji expoziční hodnotou příliš liší od zvolené výchozí expozice. Typickým příkladem takové kompozice je focení objektu proti přímému slunci, kdy dojde ke ztrátě detailů nejen ve slunečné obloze, ale také ve stinných místech scény.

Obrázek 2.5: Ilustrace rozdílu zachytitelného rozsahu okem a digitálním fotoaparátem.

Na ilustračním obrázku 2.5 lze pozorovat zjevný rozdíl mezi zachytitelným rozsahem oka a fotoaparátu. Lidskou optiku představuje černý rámeček, tu digitální rámeček červený. Pro pořízení a následnou rekonstrukci celého dynamického rozsahu scény je možné použít metodu, která dokáže skládat snímky s různou expozicí. Tato metoda je detailně popsána v rámci podkapitoly 3.3.

Kapitola 3

Obraz s vysokým dynamickým rozsahem

Využití vysokého dynamického rozsahu obrazu (zkráceně HDR[1] obrazu) nespočívá pouze v reprodukci reálného světa. Velmi aktuálním tématem je také rozsah počítačem generovaného obrazu. Generovaná grafika vytváří virtuální obraz na základě třírozměrného modelu scény. Pomocí simulace světelného záření jsou vypočteny barevné složky jednotlivých obrazových bodů ve scéně.

Klasický 8bitový přístup sice reflektuje zobrazovací schopnosti běžných monitorů, není však příliš vhodný při produkci obrazového materiálu, protože velká část původního rozsahu je nenávratně ztracena. Aby bylo možné co nepřesněji vykreslit modelovanou scénu, je nutné používat v programech pro výpočet a uchovávání hodnot datové typy s širokým rozsahem hodnot. Naivním přístupem by bylo například využití 32bitového datového typu pro každý barevný kanál. Jeden výsledný pixel tvořený třemi kanály by pak byl reprezentován 96 bity, což by znamenalo čtyřnásobnou paměťovou režii oproti klasickému 8bitovému přístupu. Tento problém řeší řada formátů pro uložení obrazu s vysokým dynamickým rozsahem, které jsou popsány v této kapitole.

3.1 Formáty pro HDR snímky

V současné době existuje celá řada formátů pro ukládání HDR obrazu, nicméně v praxi se nejčastěji lze střetnout zejména s formáty *Radiance HDR* a *OpenEXR*. Při produkci videa se tyto formáty často používají pro práci s jednotlivými snímky, které jsou až na konci pomyslného řetězce zpracování videa překódovány do výsledných video souborů.

3.1.1 Radiance HDR

Jedním z nejrozšířenějších formátů pro ukládání snímků je formát Radiance HDR, někdy též označován jako RGBE. Soubory tohoto formátu nesou příponu *.hdr. Každý soubor začíná rozšiřitelnou ASCII[2] hlavičkou. Barvy obrazových bodů jsou reprezentovány 24 bity, kde první trojice 8 bitů je vyhrazena pro mantisy jednotlivých barev, tedy pro červenou, zelenou a modrou. Posledních 8 bitů pixelu slouží pro společný exponent, který slouží jako škálovací činitel [81].

[1] HDR – vysoký dynamický rozsah (z angl. *high dynamic range*).

[2] ASCII – American Standard Code for Information Interchange: základní kódování znaků v informatice.

3.1.2 OpenEXR

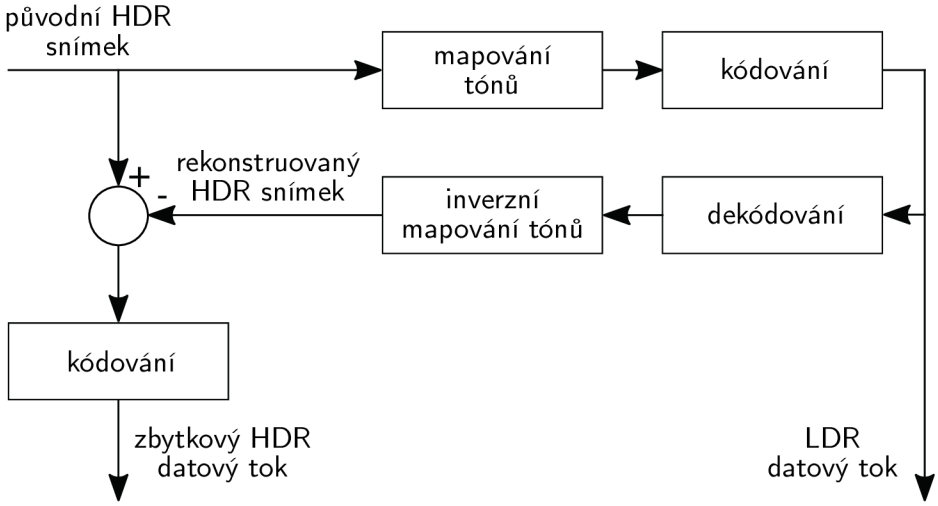
Společně s Radiance RGBE je nejpopulárnějším formátem OpenEXR[3], který vyvíjí americká firma Industrial Light & Magic[4], jež se věnuje produkci speciálních vizuálních efektů pro filmový průmysl. Soubory formátu OpenEXR mají příponu *.exr a mezi jejich klíčové vlastnosti se řadí možnost využití 16bitové a 32bitové aritmetiky s pohyblivou desetinnou čárkou a 32bitové celočíselné aritmetiky pro ukládání hodnot obrazových bodů. Formát také podporuje několik různých algoritmů pro ztrátovou i bezztrátovou kompresi. Oficiální dokumentace uvádí až 55% redukci datového objemu při použití bezztrátové komprese [42]. Lze využít volitelné množství barevných kanálu, například kanály po uložení červené, zelené a modré barvy, průhlednosti, jasové složky, hloubky a také normálových či pohybových vektorů. Je mimo jiné možné uložit obraz o několika různých rozlišeních do jednoho souboru.

Formát disponuje mnoha dalšími zajímavými vlastnostmi, které jsou detailně popsány v jeho technické specifikaci [42]. Softwarová distribuce obsahuje mnoho nástrojů pro práci se soubory a je volně dostupná ke stažení s modifikovanou svobodnou licencí, kterou lze, stejně jako samotnou distribuci, nalézt na webových stránkách tohoto projektu[3].

3.2 Formáty pro HDR video

Běžné formáty s nízkým dynamickým rozsahem využívají typicky 8 bitů pro uložení jednotlivých barevných kanálů, což odpovídá rozsahu přibližně 6 EV. Výrobci displejů klasifikují video jako HDR, jestliže jsou jednotlivé barvy zakódovány více než osmi bity. V praxi se zatím používá nejčastěji 10 bitů pro každý kanál, což umožňuje zobrazit škálu až $17,6$ EV [11].

Jedním z přístupů pro kompresi a rekonstrukci HDR videa je zpětně kompatibilní metoda. V rámci zachování možnosti zobrazení videa na běžných zařízeních s 8 bity pro každý barevný kanál byl navržen princip, díky kterému je možné zobrazit ten samý datový tok na HDR i LDR[5] displejích.



Obrázek 3.1: Schéma kódování pro reverzní rekonstrukci [31].

[3] https://openexr.com/
[4] https://ilm.com/
[5] LDR – nízký (resp. běžný) dynamický rozsah (z angl. *low dynamic range*).

Samotný princip zpětně kompatibilního přístupu není složitý. Nejprve je na výchozí HDR snímek aplikováno mapování tónů a následně proběhne kódování do zvoleného formátu v LDR podobě, který je možné zobrazit na běžných obrazovkách. Takto vzniklý obraz dále vstupuje do interního dekodéru a je na něj aplikováno inverzní mapování tónů. Vzniká tedy HDR snímek, který byl sestaven rekonstrukcí z LDR obsahu. Ten je dále odečten od prvotního snímku, čímž vzniká zbytkový obraz nesoucí data, která přesahují rámec LDR rozsahu. Schématicky je tento proces ilustrován na předchozím obrázku 3.1 [53].

Vzhledem k výpočetní náročnosti tohoto modelu se ale v praxi přistupuje k využití percepční křivky zvané elekto-optická přenosová funkce (*EOTF*[6]), která umožňuje převod dat typu `float` na celočíselné hodnoty (typicky 10bitové či 12bitové) s kvantizačním krokem, který není postřehnutelný lidským okem. Výsledné video je zakódováno do jednoho datového toku a na straně dekodéru je rekonstruováno pomocí inverzní přenosové funkce $EOTF^{-1}$. Při využití toho principu je ale nutné pro zpětnou kompatibilitu používat více nezávislých datových toků, např. jednoho pro 8bitové a druhého pro 10bitové dekodéry.

3.2.1 Standardy pro HDR video

V posledních letech vzniká celá řada doporučení pro standardy HDR videa, které vydává mezinárodní telekomunikační unie (*ITU*[7]). Aktuálně nejpoužívanější sbírka doporučení z roku 2015 nese název *ITU-R BT. 2020* a obsahuje výčet různých technických specifikací [38]. Navazuje na předchozí sadu *BT. 709*, která byla taktéž naposledy aktualizována v roce 2015 [40]. Jedná se mimo jiné o doporučení pro rozlišení obrazu, snímkovou frekvenci, bitovou hloubku, barevný prostor, reprezentaci barev a jasu, podvzorkování barev a elektro-optickou přenosovou funkci. V současné době již existuje novější sbírka doporučení *BT. 2100* aktualizovaná v roce 2018, která dále rozvíjí předchozí sadu BT. 2020 [39].

Koncový uživatel se častěji střetne s konkrétními standardy pro HDR video. V srpnu roku 2015 vydala americká organizace *CTA*[8] standard s názvem *HDR10* respektující všechna doporučení BT. 2100. Jde o volně dostupnou normu poskytovanou bez nároku na licenční poplatky, jež podporuje mimo jiné 10bitovou hloubku barevných kanálů a maximální jasovou hodnotu až 4000 cd/m^2 [19]. Jeho hlavní konkurenci představuje standard vyvíjený společností Dolby Laboratories s názvem *Dolby Vision*^{TM}, jenž byl představen v roce 2016. Mezi jeho hlavní výhody se řadí podpora až 12bitových barevných kanálů, teoretický maximální jas až 10000 cd/m^2 a možnost zápisu dynamických metadat [18, 19]. Každé zařízení podporující tento standard je ale zatíženo licenčním poplatkem. Jako odpověď na Dolby Vision^{TM} rozšířila jihokorejská společnost Samsung ve spolupráci s firmou Amazon normu HDR10 o dynamická metadata a nazvala ji *HDR10+*, přičemž byl zachován původní bezplatný koncept certifikace zařízení [50, 67].

3.2.2 Formáty kódování HDR videa

Nekomprimované video (HDR i LDR) představuje velký objem dat pro ukládání i přenos obsahu. Je tedy nutné zajistit kompresi, aby byly sníženy nároky na velikost úložišť a šíře přenosových médií. Základním požadavkem je zmenšení velikosti datového proudu při zachování co nejvyšší možné kvality vzhledem k objemu dat. Stejně jako u kódování snímků lze pro video využít ztrátový i bezztrátový přístup.

[6] EOTF – elekto-optická přenosová funkce (z angl. *Electro-Optic Transfer Function*).
[7] ITU – International Telecommunication Union.
[8] CTA – Consumer Technology Association, `https://cta.tech/`.

Pro jednodušší orientaci v této podkapitole je vhodné hned zkraje uvést základní pojmy používané v souvislosti s kompresí a kontejnerizací videa. Odborná literatura není v tomto směru příliš sjednocená, zvláště při rozlišování standardů a kompresních formátů videa. Tato diplomová práce se drží názvosloví, které striktně rozlišuje následující pojmy:

- **Standard** je předpis norem, jež je nutné splnit pro získání určité certifikace (viz 3.2.1).
  - HDR10, HDR10+, Dolby Vision™, ...
- **Formát kódování** definuje podobu dat pro přenos a ukládání souborů či toků videa.
  - H.265/HEVC, VP9, AV1, Apple ProRes, ...
- **Kontejner** svazuje různé datové toky, např. videa, audia, titulků do jednoho souboru.
  - Matroska (*.mkv), MP4 (*.mp4), QuickTime (*.mov), WebM (*.webm), ...
- **Kodek** je souhrnný název pro kódovací a dekódovací algoritmus, který zpracovává datové toky odpovídající určitému formátu kódování. Například enkodér x265 slouží pro kódování videa podle formátu H.265/HEVC. Enkodéry a dekodéry bývají často hardwarově akcelerovány, a to nejen na specializovaných čipech (např. v televizorech), ale i na novějších grafických kartách či procesorech pro běžné domácí použití. Také moderní chytré telefony podporují hardwarové dekódování pro nejběžnější formáty.
  - sady kodeků ffmpeg a libvpx, x265 (pouze enkodér pro H.265/HEVC), ...

Následuje přehled nejpoužívanějších kódovacích formátů pro HDR video. Jmenovitě se jedná o formáty H.265/HEVC, VP9 a AV1, které plně umožňují splnit všechny zmíněné standardy pro ukládání, přenos a požadovanou podobu dat videa.

MPEG H.265/HEVC

H.265/HEVC[9] je kompresní formát představený v roce 2013 skupinou MPEG[10]. Přináší úsporu až 50 % objemu dat při zachování srovnatelné kvality obrazu v porovnání se svým předchůdcem H.264/AVC[11]. Vyšší kompresní poměr je vykoupen většími výpočetními nároky pro kompresi i dekódování, nicméně s ohledem na stále rostoucí výkon dostupných zařízení tato skutečnost nepředstavuje závažný problém.

Používání formátu H.265/HEVC je stejně jako v případě jeho předchůdce zatíženo licenčními poplatky, nicméně existuje svobodný enkodér s názvem x265, jenž je financován mnoha společnostmi, které jej využívají [57]. Pro řadu firem jsou náklady spojené s licencí tak markantní, že se vydaly cestou vývoje vlastního formátu.

VP9

Přímým konkurentem pro H.265/HEVC je formát vyvinutý společností Google zvaný VP9. Firma Google, jako poskytovatel video obsahu prostřednictvím portálu YouTube[12], se rozhodla pro vývoj vlastního formátu, který zdarma poskytuje zájemcům. V roce 2010 koupila firma Google společnost On2, která vyvinula formát VP8, a rozhodla se jej následně

[9] HEVC – High Efficiency Video Coding.
[10] MPEG – Moving Picture Experts Group.
[11] AVC – Advanced Video Coding.
[12] https://youtube.com

zpřístupnit veřejnosti bez nutnosti hrazení licenčních poplatků. Formát VP8 ale nedosahoval tak kvalitních výsledků jako jeho tehdejší konkurent H.264/AVC [62], a proto přišel na řadu vývoj formátu nové generace s názvem VP9, který nabízí srovnatelné možnosti jako H.265/HEVC. Celý projekt nese název WebM[13], v jehož rámci je nabízena svobodná otevřená knihovna kodeků libvpx.

V roce 2016 vypracovala společnost Netlfix rozsáhlou studii, ve které srovnala formáty VP9 a H.265/HEVC pro několik různých rozlišení obrazu. Její závěr byl takový, že formát H.265/HEVC při použití enkodéru x265 nabízí oproti formátu VP9 kódovanému pomocí kodeku libvpx přibližně 20% úsporu dat při zachování srovnatelné kvality [60].

AV1

V roce 2015 byla založena nezisková skupina Alliance for Open Media. Zakládajícími členy byly společnosti Amazon, Cisco, Google, Intel Corporation, Microsoft, Mozilla a Netflix [1]. Později se do skupiny připojily i společnosti Apple, ARM, IBM a Nvidia, které se taktéž podílí na jejím řízení a jsou řazeny mezi zakladatele [2]. Toto společenstvo technologických gigantů si klade za cíl vyvinout svobodný formát nové generace, který bude vhodný pro přenos videa prostřednictvím internetu.

Jako základ pro formát AV1 posloužily formáty VP10 (Google), Thor (Cisco) a Daala (Xiph.org, Mozilla) [56]. První specifikace podoby datového proudu, včetně referenčního enkodéru i dekodéru, byla představena v březnu roku 2018 [74]. AV1 podporuje rozlišení až *8K*[14] při frekvenci 120 snímku za vteřinu a umožňuje použít až 12 bitů pro jednotlivé kanály. Aktuálně (listopad 2018) podporují AV1 například webové prohlížeče Firefox (Mozilla), Chrome (Google), Vivaldi a Opera, multimediální přehrávač VLC či sada kodeků FFMpeg. Platformy Youtube a Netflix v září roku 2018 spustily veřejné testování obsahu kódovaného ve formátu AV1 [59].

3.3 Pořizování HDR obrazu

V případě počítačem generované grafiky je vytváření snímků i videa poměrně přímočaré. Pořízení snímku reálného světa je ale složitější proces. Na první pohled by se mohlo zdát, že prosté zvýšení rozsahu citlivosti snímacích čipů a počtu bitů jednotlivých barevných kanálů by mohlo stačit k pořízení HDR obrazu.

Mnoho běžných digitálních fotoaparátů a kamer je schopno produkovat 12bitová, ale i 14bitová a 16bitová obrazová data v tzv. *RAW* podobě. Jedná se o data produkovaná snímačem bez dalších úprav. RAW soubor slouží jako základ pro vytvoření snímku a není vhodný pro přímé zobrazení na displeji. Uživatel má úplnou kontrolu nad transformací „surových“ dat ze senzoru na zobrazitelný snímek, jestliže se rozhodne pracovat s RAW daty. Hlavním důvodem, proč není vhodné zobrazovat přímá RAW data, je relativně velké množství obsaženého šumu a také lineární citlivost snímacích čipů. Člověk vnímá světlo s logaritmickou citlivostí, a proto, pokud fotograf nezvolí ukládání RAW dat, aplikuje aparát sérii operací s úkolem přiblížit zachycené světlo lidskému vidění a převést jej do zobrazitelného barevného prostoru. Pro tento účel bývají kamery a fotoaparáty kalibrovány pomocí tzv. *funkce odezvy* (*CRF*[15]), pomocí které mapují surová data na konkrétní hodnoty barev pixelů.

[13] https://webmproject.org/
[14] Rozlišení 8K – 7 680 pixelů pro šířku a 4 320 pro výšku videa.
[15] CRF – funkce odezvy fotoaparátu či kamery (z angl. *camera response function*).

Pro získání obrazu s opravdu vysokým dynamickým rozsahem, kde jsou barevné kanály uloženy minimálně pomocí 16 bitů, je vhodné použít techniku skládání expozic, která je popsána v následující podkapitole. Sice již existují moderní zařízení pro pořizování HDR obsahu, nicméně jde o specializované aparáty, jejichž cena je značně vysoká. Příkladem výrobců takových zařízení jsou NCTech[16] a SpheronVR[17]. Do popředí se v posledních letech dostávají společnosti Arri[18] a RED[19], které vyrábí profesionální digitální kamery pro filmový průmysl, jež podporují snímání videa s rozlišením až 8K v 16bitovém RAW režimu.

3.3.1 Skládání HDR obrazu pomocí fúze expozic

Snímek s vysokým dynamickým rozsahem je možné pořídit i prostřednictvím běžného digitálního fotoaparátu či kamery. Scénu lze nasnímat několikrát s různou rychlostí závěrky aparátu, čímž vznikne série snímků s různou expozicí. Takové snímky lze použít jako vstup pro fúzní algoritmus, který dokáže rekonstruovat původní rozsah pořizované scény.

Jako vstupní sérii je vhodné použít alespoň tři snímky. Jeden z nich je pořízen s „optimální“ expozicí, druhý je tzv. podexponovaný, tedy zachycen v kratším expozičním čase, a třetí přeexponovaný, kdy světlo na senzor dopadá oproti optimální expozici delší dobu. Existuje i dvojsnímkový přístup, kdy jsou oba vstupní obrazy odchýleny od výchozí expozice. Čím více snímků obsahuje vstupní série, tím větší dynamický rozsah je možné pokrýt. V praxi se nejčastěji lze setkat s použitím dvou až devíti snímků, protože více fotografií zvyšuje výpočetní nároky bez zásadního vlivu na výsledný rozsah a kvalitu HDR snímků.

V ideálním případě je vhodné jako vstup algoritmu používat RAW data. Je tak zajištěno, že fotoaparát na snímky neaplikoval žádné kalibrace a data zachycená snímačem jsou lineární. Lze použít i klasické fotografie ve formátu JPEG, nicméně zde je nutné zohlednit funkci odezvy, kterou aparát v rámci kalibrace dat na snímky aplikoval. Expoziční odchylky a nutný počet vstupních snímků lze stanovit empiricky na základě jejich očekávaných jasových histogramů. Platí, že pro scény s velmi vysokým dynamickým rozsahem je nutné pořídit více jednotlivých snímků s různou expozicí, aby bylo možné co nejlépe rekonstruovat jejich kompletní rozsah.

Kompoziční algoritmus

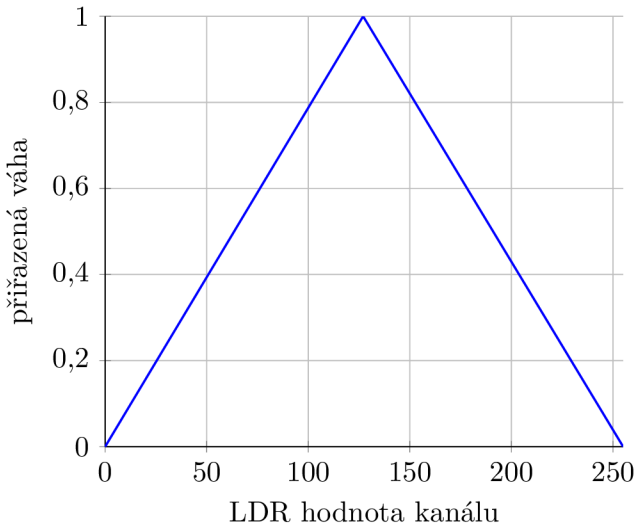
Principem fúzního algoritmu je stanovení nových hodnot pixelů ve vznikajícím HDR obrazu. Pro každý nově vznikající obrazový bod algoritmus vyčte hodnoty ze všech snímků vstupní série na dané pozici a na jejich základě stanoví hodnotu nového pixelu. Pro ilustraci následuje uvedení jednoduchého algoritmu skládání, jenž ve svém článku představili představili P. Debevec a J. Malik [16].

Tato metoda vytváří nový obraz pixel po pixelu. Hodnota každého jednoho nového obrazového bodu je vypočtena z pixelů vstupních snímků, které odpovídají souřadnicím vznikajícího bodu. Z každé fotografie ze vstupní série je tedy do výpočtu zahrnut právě jeden bod. Jsou-li jako vstup použity data ve formátu RAW, není nutné díky jejich linearitě zohledňovat funkci odezvy fotoaparátu, resp. kamery.

$$H_p = \sum_{i=1}^{N} \frac{w(L_{pi}) \cdot L_{pi})}{\Delta t_i} \Bigg/ \sum_{i=1}^{N} w(L_{pi}) \tag{3.1}$$

[16] https://nctechimaging.com/
[17] https://spheron.com/
[18] https://arri.com/
[19] https://red.com/

Rovnice 3.1 definuje výpočet nové hodnoty barevného kanálu HDR pixelu H na pozici p. N udává počet vstupních snímků, L reprezentuje jeden pixel původního LDR bodu na pozici p snímku i. Expoziční čas snímku i je vyjádřen jako Δt. Do rovnice vstupuje také váhová funkce w, která danému bodu pro všechny vstupní snímky přiřadí určitou váhu. Debevec a Malik ve svém článku navrhují velmi jednoduchou funkci, kterou lze matematicky vyjádřit následujícím grafem.



Váhovou funkci je možné upravit např. na obdobu Gaussovy funkce [66], nebo ji dokonce adaptivně modifikovat pro každý vstupní snímek [43]. Existuje velké množství možností, jakým způsobem lze nakládat se vstupní sekvencí a vytvářet HDR obraz.

3.3.2 Pořizování vstupních snímků pro fúzní algoritmy

Fotografická technika, kdy je scéna nasnímána do několika různých expozic, se nazývá *expoziční bracketing*. Jediný parametr fotoaparátu, který je vhodné měnit pro dosažení různých expozic, je rychlost závěrky, protože změna nastavení clony mění hloubku ostrosti fotografií a rozdílná citlivost ISO způsobí nerovnoměrný výskyt šumu na jednotlivých snímcích. Metody pro docílení expozičního bracketingu lze rozdělit na dvě kategorie:

Sériové snímání scény

Sériové (též sekvenční či temporální) snímání je nejjednodušší metoda pro vytváření vstupních sérií snímků, protože k realizaci nejsou nutná specializovaná zařízení. Uživatel si vystačí pouze s jedním běžným digitálním fotoaparátem nebo kamerou, jelikož snímky jsou pořizovány sekvenčně jeden po druhém. Nutností je zajistit co nejmenší pohyb aparátu i objektů ve scéně v průběhu snímání.

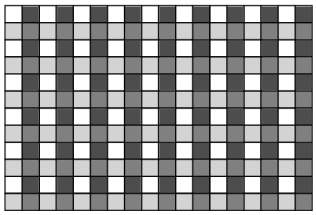
První možností pro sériové snímání je ruční nastavení rychlosti závěrky pro jednotlivé fotografie. Velkou nevýhodou manuálního nastavení je čas, který uplyne mezi zachycením prvního a posledního snímku. Pokud se během pořizování změní snímaná scéna, bude výsledný HDR snímek obsahovat pohybové artefakty. Vzhledem k tomu, že algoritmy skládání pracují typicky pixel po pixelu a nerespektují obsah scény, představuje i pohyb samotného snímače problém. Výskyt pohybových artefaktů lze částečně eliminovat dodatečně pomocí algoritmů pro zarovnání snímků a detekci pohybu objektů ve scéně [78].

Běžné moderní aparáty podporují tzv. *softwarový autobracketing*, který umožňuje uživateli zadat parametry snímání a následně provede sekvenční snímání jedním stisknutím spouště. Samotné pořízení všech snímků tak proběhne výrazně rychleji, než při manuálním snímaní jedné fotografie po druhé. Princip autobracketingu je používán i na pozadí HDR režimu implementovaného v běžných digitálních fotoaparátech a chytrých telefonech [3]. Výstupem takových zařízení nebývá opravdový HDR obraz, ale obraz s běžným dynamickým rozsahem získaný pomocí mapování tónů, což platí pro statické snímky i pro video.

Paralelní snímání scény

O paralelním přístupu lze hovořit, jestliže pořizování všech LDR snímků probíhá zároveň. Existuje několik způsobů, jak docílit paralelního snímání:

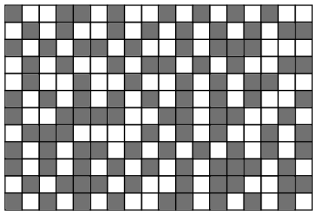
- **Použití neutrálních filtrů**
  Tento přístup využívá optické *atenuace*[20] k vytvoření HDR obrazu. Typicky je plocha čipu pokryta neutrálními šedými filtry, které jsou uspořádány do pravidelné mřížky [58]. Pro snazší pochopení je taková mřížka filtrů schématicky vyobrazena na obrázku 3.2.



Obrázek 3.2: Pravidelná mřížka filtrů na snímači.

  Na předchozím schématu je ilustrováno použití čtyř různě propustných filtrů. V jednom okamžiku na snímači vznikají simultánně 4 snímky s rozdílnou expozicí, které jsou na čipu vzájemně prolnuté. Nevýhodou tohoto přístupu je pokles efektivního rozlišení snímače. V ilustrovaném případě vzniká HDR obraz se čtvrtinovým rozlišením oproti plnému rozlišovacímu potenciálu snímacího čipu.

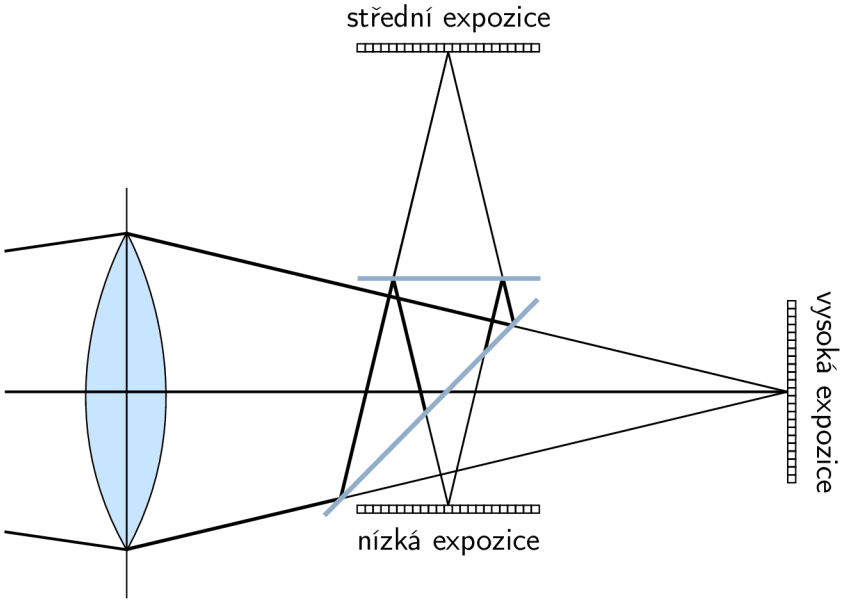
  Existuje i jednočipová metoda snímaní umožňující pořídit snímek v plném rozlišení snímače. Princip spočívá v nerovnoměrném uspořádání neutrálních filtrů. Příklad takového přístupu s použitím jednoho druhu filtrů je ilustrován na obrázku 3.3. Rekonstrukce obrazu je možná díky frekvenčně selektivní extrapolaci [70, 71]. Kvalita a věrnost výsledného obrazu sice nejsou tak vysoké jako u ostatních metod snímaní, nicméně pro některé případy použití může být tato metoda vhodná.



Obrázek 3.3: Nepravidelná mřížka filtrů.

[20] Optická atenuace – utlumení intenzity světla pomocí průchodu určitým materiálem.

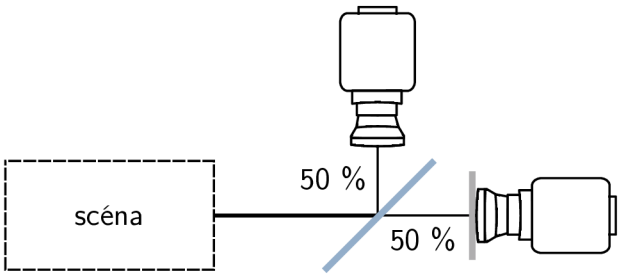
- **Použití několika snímacích čipů**
  Rozdílných expozic lze docílit i použitím několika snímacích čipů v rámci jednoho zařízení. Paprsky procházející objektivem jsou pomocí optických prvků částečně pohlceny a usměrněny na jednotlivé snímače. Příklad takového přístupu je uveden na obrázku 3.4 [76].



Obrázek 3.4: Více snímačů v jednom zařízení [76].

Závěrka je otevřena delší dobu oproti střední expozici a paprsky přímo dopadají na senzor, kde vzniká obraz s nejvyšší intenzitou jasu, jenž je na schématu popsán jako „vysoká expozice“. Světelný proud potkává cestou k tomuto senzoru dva rozdělovače paprsků, které část světla odrazí a část propustí. Na zbylé dva snímače tak dopadá světlo s nižšími intenzitami jasu oproti senzoru prvnímu.

- **Použití několika snímacích zařízení**
  Velmi podobného principu, jako v případě použití několika snímacích čipů, lze dosáhnout při snímaní scény několika aparáty. Schéma 3.5 ukazuje jednu z možností použití dvou zařízení k zachycení scény. Paprsky odražené od objektů ve scéně jsou pomocí polopropustného zrcadla rovnoměrně rozděleny a před jednou z čoček kamer je umístěn neutrální šedý filtr [32]. Současně tak kamery zachytávají dva obrazy s rozdílnou expozicí.



Obrázek 3.5: Snímaní scény několika kamerami.

Kapitola 4

Mapování tónů

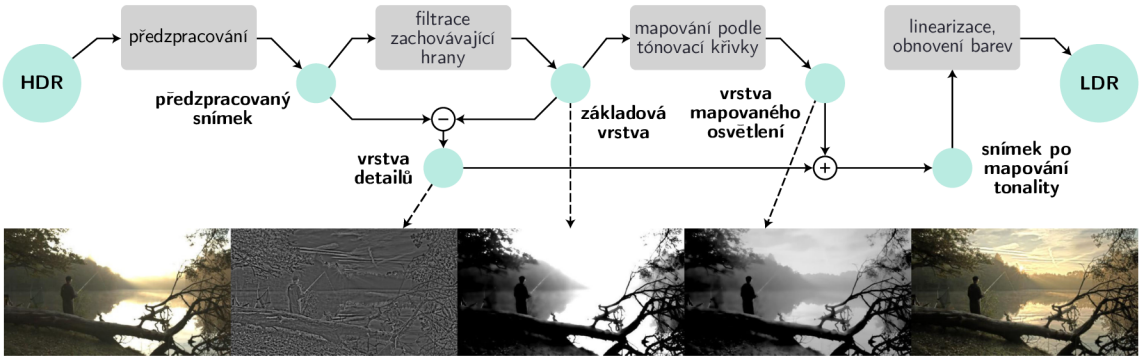
Pomocí metod popsaných v sekci 3.3 lze zachytit celý dynamický rozsah pořizované scény. Dnešní displeje takto vysoký rozsah nepodporují a vzniká tedy nesoulad, kdy sice existují HDR fotografie a video, ale není možné na běžných displejích nativně zobrazit jejich obsah. Tento problém řeší metoda zvaná *mapování tónů*, která zajišťuje převod HDR obrazu na reprezentaci se zobrazitelným rozsahem.

V principu je tedy mapování tónů algoritmus pro kompresi dynamického rozsahu. Pro jednotlivé algoritmické metody mapování tónů používá drtivá většina odborné literatury označení *operátory mapování tónů*, jež je často redukováno na zkratku *TMO* (z angl. *tone mapping operator*). Samotné mapování je prováděno pomocí určité matematické funkce.

Prostá lineární komprese rozsahu nepřináší výsledky, které by vizuálně odpovídaly původní scéně, a je nutné použít pro převod složitější křivku, například logaritmickou, exponenciální či logistickou. Některé operátory využívají funkci odvozenou od jasového histogramu obrazu, nebo od jeho gradientní domény [24]. Nejjednodušší kategorizací operátorů tónového mapování je rozlišování tzv. *globálních* a *lokálních* operátorů.

- **Globální operátory** nejprve analyzují snímek jako celek a následně aplikují na každý obrazový bod stejnou funkci pro mapování tónů.

- **Lokální operátory** zohledňují při výpočtu nových hodnot i okolí daného bodu. Ve srovnání s globálními metodami je typicky zajištěno lepší zachování detailů ve všech oblastech obrazu. Obecný princip fungování lokálních algoritmů lze schématicky rozdělit na několik klíčových fází, které jsou vyobrazeny na následujícím obrázku 4.1.



Obrázek 4.1: Schéma typického algoritmu lokálního mapování tónů [24].

V prvním kroku lokálního mapování je provedena série operací, která předzpracuje obraz pro další fáze. Typicky může jít o škálování, extrakci relativní luminance, či transformaci snímku do logaritmické domény. Následuje filtrace zachovávající hrany objektů, díky které vzniká tzv. *základová vrstva*, jež neobsahuje detaily a nese pouze informace o osvětlení objektů ve scéně. Pro tento účel slouží např. bilaterální, nebo řízený filtr. Dalším krokem je extrakce tzv. *detaily nesoucí vrstvy*, jež vznikne odečtením základové vrstvy od původního předzpracovaného snímku. Na základovou vrstvu je poté aplikována mapovací funkce, označována jako *tónovací křivka*, a následně je k ní přičtena vrstva s detaily. V poslední fázi proběhne obnovení barev, linearizace obrazu, a případně další finální operace nad výsledným obrazem se zobrazitelným dynamickým rozsahem.

4.1 Extrakce základové a detaily nesoucí vrstvy

Než tato práce přistoupí k popisu jednotlivých operátorů mapování tónů, je nutné podrobně vysvětlit samotné dílčí kroky lokálního mapování tónů. Protože předzpracování obrazu je většinou unikátní pro každý operátor, nebude se mu tato část textu věnovat. Prvním obecným krokem, jehož princip je pro velkou většinu lokálních operátorů velmi podobný, je filtrace sloužící pro extrakci základové a detaily nesoucí vrstvy.

Z pohledu fyzikální optiky popisuje základová vrstva intenzitu osvětlení objektů ve scéně. V případě vrstvy detailů pak lze hovořit o *reflektanci* neboli o *odrazivosti* jejich povrchů. Pozorovatel je schopen díky adaptaci svého oka rozlišovat odrazivost objektů ve scéně ve velkém rozsahu a z pohledu mapování tónů lze reflektanci objektů vnímat jako invariantní vůči jejich osvětlení [35]. Operátory tedy mohou využít původní odrazivost objektů a komprimovat pouze celkové osvětlení, díky čemuž je zajištěno zachování lokálního kontrastu [24].

Následuje popis vybraných metod filtrace pro extrakci základové vrstvy osvětlení, která je dále použita pro získání vrstvy reflektance objektů nesoucí lokální detaily (viz schéma 4.1). Ta je získána jako zbytkový obraz po odečtení základní vrstvy osvětlení od původního obrazu s vysokým dynamickým rozsahem.

4.1.1 Gaussův filtr

Gaussův filtr (též Gaussián) představuje výpočetně jednoduchou metodu pro aproximaci osvětlení scény. Principem tohoto filtru je výpočet nové hodnoty bodu na základě váženého průměru hodnoty daného pixelu a jeho okolí. Váhy okolních bodů jsou určeny na základě Gaussova rozložení, kdy pixely, které jsou z hlediska prostorové vzdálenosti blíže centrálnímu bodu, mají vyšší váhu než body vzdálenější. Největším problémem tohoto konvolučního lineárního filtru je velmi neostrý přechod výsledné intenzity v oblastech hran objektů. Ve vznikajícím obraze se pak tato skutečnost vizuálně projevuje jako „záře“ jednotlivých objektů. Tento artefakt zahraniční literatura označuje jako tzv. *halo efekt*. Konvoluční jádro Gaussova filtru lze vyjádřit rovnicí 4.1 [75].

$$GF(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{4.1}$$

Stupeň vyhlazení obrazu je dán parametrem rozptylu σ. Čím větší je hodnota σ, tím větším okolím je hodnota konkrétního bodu ovlivněna, což má za následek větší vyhlazení jasových přechodů. Vzdálenost bodu na horizontální, resp. vertikální, ose od počátku filtru udává x, resp. y.

Gaussův filtr není příliš vhodný pro účely mapování tónů, protože při výpočtu nedokáže zohlednit hrany objektů. V tomto textu je uveden zejména z ilustračních důvodu a také proto, že slouží jako pomyslný stavební kámen pro pokročilejší metody filtrace zachovávající hrany objektů. Často používaným filtrem pro extrakci základové a detaily nesoucí vrstvy pro účely lokálního mapování tónů je například bilaterální filtr.

4.1.2 Bilaterální filtr

Tato metoda byla popularizována článkem *Bilateral Filtering for Gray and Color Images* z roku 1998 [77]. Podobně jako Gaussián je bilaterální filtr definován jako vážený průměr bodu a jeho okolí. Hlavní rozdíl spočívá v tom, že bilaterální filtr při výpočtu vah zohledňuje i samotné hodnoty okolních pixelů. Hlavní myšlenkou tedy není jen aplikovat na skupinu bodů průměrující funkci, ale také aktivně ohodnotit a uvážit jejich podobnost. Díky tomu filtr zachovává ostrost hran a dosahuje výrazně lepších výsledků oproti Gaussově filtru v kontextu mapování tónů. Matematicky lze rovnici obecného bilaterálního filtru zapsat jako [5, 23, 72]:

$$BF_s = \frac{1}{k(s)} \sum_{p \in \Omega} f(p - s) \cdot g(I_p - I_s) \cdot I_p. \tag{4.2}$$

Hodnota p z rovnice 4.2 představuje pozici centrálního pixelu, proměnná s značí souřadnice konkrétního zpracovávaného bodu z okolí Ω. I_p, I_s vyjadřují hodnoty bodů o souřadnicích p, resp. s. Pro základní implementaci klasického bilaterálního filtru je typické využití dvou Gaussových funkcí f, g:

$$f(x) = g(x) = e^{-\frac{x^2}{2\sigma^2}}. \tag{4.3}$$

Tyto funkce jsou na sobě nezávislé a hodnotu rozptylového parametru σ lze definovat pro každou z dvojice funkcí zvlášť. Přiřazení vah hodnotám pixelů podle prostorové vzdálenosti vůči centrálnímu bodu zajišťuje funkce f, zatímco funkce g přiřazuje okolním bodům váhu podle podobnosti jejich hodnoty, taktéž vůči středovému pixelu. Body s podobnými vlastnostmi jsou ohodnoceny vyšší váhou než body opticky rozdílné. V původní rovnici 4.2 figuruje také normalizační faktor $k(s)$, který lze vyjádřit následovně:

$$k(s) = \sum_{p \in \Omega} f(p - s) \cdot g(I_p - I_s). \tag{4.4}$$

Existuje mnoho variant tohoto nelineárního filtru, které se vzájemně liší použitými funkcemi pro výpočet váhy, výpočetní rychlostí, či počtem průchodů obrazem. Metoda bilaterálního filtrování nachází uplatnění mimo mapování tónů i v dalších oblastech zpracování obrazu, například při redukci obrazového šumu.

4.1.3 Řízený filtr

Hlavní motivací pro vznik řízeného filtru (angl. *guided filter* [36]) byla zejména velká výpočetní náročnost a také nedostatečná přesnost bilaterálního filtru. Časová složitost řízeného filtru je vždy lineární a není závislá na velikosti jeho konvolučního jádra. Oproti rychlosti vyniká vůči bilaterálnímu filtru i chováním v oblastech hran, protože netrpí v takové míře artefakty způsobenými převrácením gradientu obrazu okolo hran objektů. Takový artefakt se na výsledném snímku může projevovat jako nežádoucí barevná záře jednotlivých objektů ve scéně.

Předpokladem řízeného filtru je existence lineárního modelu mezi vstupním řídícím snímkem I a výstupním snímkem q. Tuto skutečnost lze lokálně aplikovat na určitou oblast snímku označenou jako ω_k se středem na pozici k:

$$q_i = a_k I_i + b_k, \forall i \in \omega_k, \tag{4.5}$$

kde a_k, b_k jsou lineární koeficienty, u nichž lze předpokládat jejich konstantní hodnoty v rámci lokálního „okna“ ω_k. Proměnná i reprezentuje pozici zpracovávaného pixelu. Tento lokálně lineární model zajišťuje, že hrana se na výstupním snímku q vyskytne pouze tehdy, jestliže existuje v řídícím snímku I, protože $\nabla q = a \nabla I$, kde ∇ vyjadřuje gradient obrazu. Výpočet koeficientů a_k, b_k spočívá v nalezení takového řešení rovnice 4.5, které minimalizuje rozdíl mezi výstupem q a vstupem p. Konkrétně se jedná o nalezení nejmenšího řešení ztrátové funkce:

$$E(a_k, b_k) = \sum_{i \in \omega_k} ((a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2), \tag{4.6}$$

kde ε představuje regularizační parametr snižující vliv velkých hodnot koeficientu a_k. Ztrátová funkce (*angl. loss function*) E může nabývat hodnoty od nuly do jedné. Čím je hodnota nižší, tím je výsledek lepší. Pro nalezení co nejmenšího možného řešení rovnice 4.6 slouží tzv. lineární regrese:

$$a_k = \frac{\frac{1}{|\omega_k|} \sum_{i \in \omega_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon}, \quad b_k = \bar{p}_k - a_k \mu_k. \tag{4.7}$$

Střední hodnotu a odchylku řídícího obrazu I v rámci okna ω_k vyjadřují μ_k, resp. σ_k^2. Počet obrazových bodů v okně udává $|\omega_k|$ a jejich aritmetický průměr $\bar{p}_k$ lze vypočítat pomocí rovnice 4.8:

$$\bar{p}_k = \frac{1}{|\omega_k|} \sum_{i \in \omega_k} p_i. \tag{4.8}$$

V další fázi probíhá aplikace lineárního modelu na všechna lokální okna snímku. Výsledné obrazové body na pozici i mohou pro různá okna, ve kterých se nachází, nabývat různých hodnot. Finální výsledná hodnota obrazového bodu q_i je pak dána průměrem všech takto získaných hodnot. Po získání koeficientů a_k, b_k pro všechna obrazová okna ω_k tedy lze vypočítat výsledek filtru pomocí následující rovnice:

$$q_i = \frac{1}{|\omega_k|} \sum_{k: i \in \omega_k} (a_k I_i + b_k) = \bar{a}_i I_i + \bar{b}_i, \tag{4.9}$$

kde průměrné hodnoty $\bar{a}_i$ $\bar{b}_i$ lze vyjádřit pomocí rovnic:

$$\bar{a}_i = \frac{1}{|\omega_k|} \sum_{k \in \omega_i} a_k, \quad \bar{b}_i = \frac{1}{|\omega_k|} \sum_{k \in \omega_i} b_k. \tag{4.10}$$

Po provedení předchozí modifikace už nadále neplatí linearita mezi gradienty ∇q a ∇I, protože lineární koeficienty $\bar{a}_i, \bar{b}_i$ se prostorově liší. Lze ale předpokládat, že pokud jsou $\bar{a}_i, \bar{b}_i$ výstupy průměrování, budou jejich gradienty v okolí ostrých hran objektů signifikantně menší než v případě řídícího obrazu I a platí tedy $\nabla q \approx \bar{a} \nabla I$. Díky této skutečnosti je zaručeno, že rapidní změny hodnot intenzit řídícího obrazu I jsou ve výstupním obrazu q zachovány.

4.2 Mapování rozsahu podle tónovací křivky

Při pohledu na obecný princip samotného mapování tónů je zřejmé, že jde ve své podstatě o matematickou funkci, která mapuje hodnoty z jednoho rozsahu na hodnoty z rozsahu druhého. V rámci mapování tónů lze takovou funkci označit jako tónovou či tónovací křivku.

Základním cílem operátoru mapování tónů je přiřazení HDR snímkům z teoretického rozsahu $L \in \mathbb{R}^+$ hodnoty, které lze zobrazit na běžném displeji, tedy $l \in \mathbb{N} = \{0, \ldots, 255\}$ pro červenou, zelenou a modrou barvu. Tento proces lze vyjádřit jako obecnou transformaci $M : L \rightarrow l$ [24]. Požadavkem na algoritmus je zachování globální lineární podobnosti výsledného obrazu vůči vstupním HDR datům. Různé operátory řeší tento požadavek rozdílnými metodami. Zatímco některé mapují data přímo na hodnoty pixelů $\{0, \ldots, 255\}$, jiné mapují vstupní rozsah například na svítivost konkrétního zobrazovacího zařízení a potřebují pro výpočet parametrizovaný model displeje. V tomto kontextu lze popsat tónovací křivku jako $V : L \rightarrow T$, kde $T \in \mathbb{R}^+$, přičemž následuje finální transformace na konkrétní hodnoty zobrazitelné na daném zařízení [25, 54].

4.2.1 Lokální adaptace tónovací křivky

Tónovací křivka bývá zpravidla definována tak, že při kompresi rozsahu dochází i ke ztrátové kompresi kontrastu. Zatímco globální kontrast musí být nevyhnutelně redukován, aby bylo možné dodržet výsledný rozsah, lokální kontrast, a tím pádem i lokální detaily, je možné vůči výstupnímu globálnímu kontrastu zesílit.

Lokální operátory mapování tónů typicky rozdělují obraz na tzv. *základovou vrstvu* (BP) a *detaily nesoucí vrstvu* (DP), která vznikne odečtením základové vrstvy BP od původního HDR snímku (viz podkapitolu 4.1). Tónovací křivka je poté lokálně aplikována pouze pro základovou vrstvu BP a vrstva detailů je k výslednému snímku přičtena beze změn. Celý proces lokální adaptace bývá zpravidla vykonáván v logaritmické doméně [24].

Podoba tónovací křivky typicky závisí na sadě parametrů specifikovaných uživatelem a na statistikách získaných z původních HDR dat. Příkladem omezení může být například maximální jas displeje, či jeho kontrastní vlastnosti. Typickou statistikou získanou z původního HDR obrazu je průměrná hodnota jasu, maximální a minimální hodnota jasu nebo histogram snímku. Pro lepší zachování lokálních detailů ve všech částech scény lze pro jednotlivé oblasti obrazu zohlednit statistické hodnoty získané z lokálního okolí právě zpracovávaného bodu. Při použití tohoto přístupu je možné adaptovat tónovou křivku podle okolí, ve kterém se daný pixel nachází. Lokální statistiky lze přímo použít při stanovení podoby tónovací křivky, která tak může mít různou podobu pro rozdílné oblasti snímku [24].

4.3 Rozdíly mezi mapováním tónů statických snímků a videa

Mapování tónů napříč sekvencí snímků přináší oproti zpracování jednotlivých statických obrázků mnoho dalších problémů. Největším rozdílem je bezesporu požadavek na dosažení temporální koherence mezi navazujícími snímky. Bez uvážení této skutečnosti dochází ve výsledném videu ke vzniku nežádoucích obrazových artefaktů. Druhým požadavkem je zachování globálního poměru kontrastu jak pro obsah jednotlivých snímků, tak pro celé video. Dalším problémem, který musí být zohledněn, je výpočetní náročnost mapovacího algoritmu, protože je často nutné operátory mapování tónů spouštět pro velké množství vstupních snímků, nebo je aplikovat v reálném čase.

4.3.1 Obrazové artefakty

Naivní přístup pro temporální mapování tónů lze popsat tak, že mapování probíhá pro každý vstupní snímek zvlášť, bez ohledu na okolní snímky ze sekvence. Takový postup vede k výskytu obrazových artefaktů ve výsledném videu, které jsou způsobeny změnami napříč původními snímky v temporální doméně. Existuje mnoho různých druhů artefaktů, které lze v kontextu temporálního mapování tónů definovat. Popis artefaktů v tomto textu byl inspirován článkem *Survey of Temporal Artifacts in Video Tone Mapping* z roku 2014 [8].

Jasové artefakty

První kategorii temporálních artefaktů lze obecně charakterizovat jako problém jasové koherence po sobě jdoucích snímků. Globální problikávání (angl. *flickering*) celého obsahu jednotlivých snímků se projevuje v případě, že operátor mapování tónu mění nepřiměřeně své parametry mezi jednotlivými snímky, nebo jestliže ve vstupní HDR sekvenci po sobě následují snímky s velkým rozdílem maximální či minimální hodnoty globálního jasu. Díky této skutečnosti hrozí, že oblasti, které si napříč snímky jasově odpovídají, budou transformovány do jiného rozsahu výsledné intenzity. I malá změna osvětlení scény tak může způsobit skokovou změnu celkového jasu ve výsledném videu. Následující sekvence obrázků 4.2 ilustruje tento druh artefaktu na testovacím videu *Campus Norrköping Hallway*[1] [26].

Obrázek 4.2: Příklad globálního problikávání neboli flickeringu. Druhý snímek v sekvenci disponuje velmi rozdílnými jasovými vlastnostmi oproti ostatním snímkům z výsledné sekvence, což se v konečném videu projeví jako bliknutí.

Problikávání může nastat i lokálně, pro menší oblasti snímků. Typicky se projevuje pouze u lokálně pracujících operátorů mapování tónů. Princip je ve své podstatě stejný jako u globálního problikávání, ovšem v tomto případě dochází ke ztrátě jasové koherence pouze v určité oblasti snímku.

Artefakty spojené s kontrastem

K problémům z hlediska kontrastu může docházet, jestliže operátor mapování tónů reaguje na změnu osvětlení scény posunem rozsahu kontrastu, kterým kompenzuje změnu osvětlení. Jasově se tedy scéna muže jevit jako konzistentní, nicméně kontrast, jenž by odpovídal původní scéně, není zachován, protože operátor nereaguje na alteraci osvětlení posunem celkového jasu obrazu.

Může také docházet k problémům z hlediska koherence kontrastu jednotlivých objektů ve scéně. Projevem je pak změna kontrastu i jasu v oblastech, které v původních snímcích s vysokým dynamickým rozsahem zůstávají v temporální doméně stabilní. Tento artefakt je do jisté míry akceptovatelný, protože lidské oči taktéž provádí adaptaci vidění na základě intenzity osvětlení konkrétní scény.

[1] Linköpings Universitet, repozitář HDR videí: http://www.hdrv.org/Resources.php.

Obrázek 4.3: Po provedení mapování tónů dochází k nesprávné temporální interpretaci kontrastu v oblasti mostu, jehož vizuální vlastnosti nejsou vůči reálné scéně zachovány. Sekvence snímků pochází z volně dostupné databáze HDR videí[2] Ronana Boitarda [8, 9].

Pohybové artefakty

V českém fotografickém názvosloví se pro tento problém typicky používá pojem „duchové“, ale často se lze setkat i s pojmem *ghosting*, který byl přejat z anglické odborné literatury. K tvorbě duchů může docházet ve dvou různých fázích produkce HDR videa. Prvním případem je zanesení duchů do obrazu již v průběhu pořizování HDR snímků. Zejména metoda fúze sekvenčně pořízených snímků z rozdílnou expozicí, které se tento text věnuje v části 3.3.1, může produkovat duchy na HDR snímcích. Ti mohou vznikat v důsledku pohybujícího se objektu ve scéně nebo pohybem samotného snímacího zařízení během pořizování snímků.

Obrázek 4.4: Po složení prvních třech snímků s rozdílnou expozicí vznikají na výsledném čtvrtém snímku vlivem pohybu plamenů svíček duchové. Obrázky pochází z datové sady[3] dodané k článku, jenž popisuje algoritmy pro redukci ghostingu [80].

Jestliže na scéně dojde k pohybu objektů, při kombinaci expozic se pohybující objekty objevují na několika místech výsledného obrazu s jistou úrovní transparence. To je důvod, proč se tento druh artefaktu označuje právě jako „duchové“. Fúzní algoritmus typicky prochází snímky bod po bodu a nezahrnuje korekci pohybu. Detekci a odstranění duchů neboli *deghostingu*, se věnuje velké množství specializovaných algoritmů [80].

Druhu příčinou vzniku pohybových artefaktů může být samotný operátor mapování tónů. Obzvláště operátory pracující v gradientní doméně obrazu jsou na tento druh artefaktu velmi citlivé [29, 47, 48]. Nejjednodušší přístup, který uvažuje pohyb ve scéně, počítá odhad pohybu pro jednotlivé obrazové body HDR snímků. Pokročilejší metody používají typicky konvoluční jádro pro identifikaci objektů či jejich ucelených částí, pro které následně stanovují křivky pohybu (angl. *motion paths*). Je-li odhad pohybu dostatečně spolehlivý, lze provést filtraci v čase jak v neměnných, tak v pohybujících se částech obrazu. Problémy do tohoto konceptu ovšem vnáší překrývání objektů ve scéně, kdy je stanovení pohybových křivek velmi problematické [24].

[2] Databáze HDR videí Ronana Boitarda: http://www.ece.ubc.ca/~rboitard/.
[3] Datová sada [80]: http://user.ceng.metu.edu.tr/~akyuz/files/eg2015/index.html.

4.3.2 Temporální koherence videa

Hned z kraje této části textu je vhodné jasně vymezit pojmy *temporální adaptace* a *temporální koherence*, na které se tento text v následujících kapitolách často odkazuje:

- **Míra temporální adaptace** vyjadřuje, jak efektivně dokáže operátor mapování tónů využít původní dynamický rozsah jednotlivých vstupních snímků.
- **Míra temporální koherence** udává, zda vizuální vlastnosti výstupního videa odpovídají vlastnostem původní sekvence a s jakou mírou se na něm projevují nežádoucí temporální artefakty.

Požadovanou vlastností temporálních operátorů je zajištění co nejvyšší možné míry adaptace i výsledné koherence. Cílem je tedy vytvořit temporálně koherentní video a zároveň využít co největší potenciál dynamického rozsahu všech snímků ze vstupní sekvence. Tyto dva požadavky jsou ovšem protichůdné a často je nutné obětovat míru adaptace, aby výstup zůstal koherentní. Není-li algoritmus schopen „vhodně“ modifikovat své výpočty v temporální doméně, dochází k problémům koherence výsledného videa v podobě temporálních artefaktů.

Z hlediska lidské zrakové soustavy lze jako hlavní nežádoucí artefakty označit problikávání a ghosting, popsané v předchozí části textu. Artefakty spojené s kontrastem jsou pro lidské oči přirozenější, protože zrak se v reálném světě také adaptuje na různé světelné podmínky. Kontrastní artefakty tedy lze do určité míry akceptovat, zvláště pokud se ve výsledném videu neprojevují skokově, či pokud nedochází k vysoké míře divergence vůči vstupním datům [8].

Mnoho obrazových statistik, které jsou zahrnuty při výpočtu tónovací křivky, se může rapidně změnit s přechodem na další snímek ze sekvence, což má za následek nežádoucí skokovou změnu tónovací křivky. Taková změna se pak projeví problikáváním. V případě globálních operátorů, jež používají pro všechny obrazové body stejnou tónovací křivku, lze docílit hladkého přechodu mezi jednotlivými snímky použitím filtru typu dolní propust, který při výpočtu statistik uvažuje předešlé i následující snímky vůči snímku právě zpracovávanému. Činnost takového filtru ovšem může být negativně ovlivňována snímky, které ve zpracovávané sekvenci následují po snímku aktuálním. Přesněji řečeno, operátor může začít reagovat na změny ve scéně dřív, než k nim doopravdy dojde. Přirozeným řešením zmíněného problému může být použití filtru s nekonečnou impulzní odezvou prvního řádu, který zohledňuje hodnoty statistik pouze ze snímků předešlých [24]:

$$p_s(L_t) = \alpha \cdot p_s(L_t) + (1 - \alpha) \cdot p_s(L_{t-1}). \tag{4.11}$$

Proměnná p_s z rovnice 4.11 představuje obecný statistický parametr, α představuje impulzní odezvu filtru a L_t je podoba tónovací křivky v diskrétním čase t.

Další metodou pro dosažení temporální koherence může být dodatečné zpracování výstupních snímků, které může být aplikováno na libovolný operátor mapování tónů. V tomto ohledu lze samotný operátor vnímat jako černou skříňku, která produkuje vstup pro dodatečnou úpravu. Cílem je dosažení temporální koherence za použití určitých výpočetních operací pro odstranění artefaktů. Problémem je, že algoritmy vyžívající podobných principů často pracují jen na globální úrovni, kde aplikují škálování jasu tak, aby nedocházelo k viditelnému problikávání. Existuje i několik lokálních metod pro dodatečné zpracování, nicméně jejich hlavní nedostatek spočívá v jejich specializaci, protože jsou typicky navrženy pro redukci pouze jednoho druhu, nebo omezené množiny artefaktů [10, 20, 46].

4.3.3 Viditelnost a zesílení šumu

Jednou z důležitých implikací procesu pořizování HDR videa je zvýšení množství šumu na snímcích, jmenovitě v tmavých oblastech scény. Zejména při použití kamer pořizujících záznam s vysokou bitovou hloubkou je šum velkým problémem, zvláště je-li nutné zachytit tmavou scénu v krátkém expozičním čase, kdy dochází k zesílení zachycených dat snímacím čipem. Tímto problémem se text zabývá v podkapitole 2.2.2. I přestože šum nemusí být na původních datech zřetelný, po aplikaci lokálního mapování tónů může být situace zcela odlišná. Vzhledem k tomu, že většina operátorů nepoužívá pro zesílení jasu a kontrastu v tmavých oblastech scény lineární tónovací křivku, dochází i k významnému zesílení šumu v těchto částech obrazu [24].

Nabízí se zde možnost použití dodatečného algoritmu pro odstranění šumu, což je ovšem výpočetně náročné řešení. Některé operátory redukují šum jako vedlejší efekt jejich činnosti, pakliže pracují globálně a filtrování probíhá pro každý pixel zvlášť. Pro lokální mapování tónů je v principu nemožné odstranit všechen šum bez vytváření nových nežádoucích artefaktů. I přesto vznikají za účelem zvýšení robustnosti vůči obrazovému šumu lokální operátory, které při vytváření tónovací křivky explicitně počítají charakteristiku šumu [25] pro dané sekvence. Tato stanovená charakteristika slouží pro detekci a redukci šumu v obrazu.

4.4 Přehled vybraných operátorů pro mapování tónů videa

Tato podkapitola nabízí přehled vybraných operátorů pro mapování tónů HDR videa. Zvoleny byly ty operátory, jež jsou historicky významné, a ty, které článek *A comparative review of tone-mapping algorithms for high dynamic range video* [24] popisuje jako nejzajímavější. Uvedený seznam je řazen chronologicky a každý z uvedených operátorů se vyznačuje unikátním přístupem pro mapování tónů videa. V této části textu je často používána zkratka *TMO*, která představuje „operátor mapování tónů“ (z angl. *tone mapping operator*).

- **Visual Adaptation TMO** [30], publikovaný v roce 1996, je založený na psychofyzikálních experimentech, které popisují vztahy mezi jasovými a barevnými vlastnostmi obrazu a jejich vlivem na lidský zrak. Experimenty uvažují unikátní percepční vlastnosti tyčinek i čípků lidské sítnice. Model popsaný těmito vztahy slouží pro výpočet míry adaptace aplikované na obraz s předem stanovenou základní hodnotou jasu.

- **Interactive Tone Mapping** [22] je založen na operátoru popsaném v předchozím bodě a ve svém algoritmu používá dva průchody napříč sekvencí snímků. Při prvním průchodu algoritmus prochází podvzorkované snímky se sníženým rozlišením v logaritmické doméně a dochází k výpočtu míry adaptace. Vypočtené hodnoty následně aplikuje v druhém průchodu v podobě parametrů pro samotné mapování tonality jednotlivých snímků.

- **Mal-adaptation TMO** [41] si klade za cíl reprodukovat obraz co nejvěrněji vůči vnímaní scény lidským zrakem, který může být jak zdravý, tak nějakým způsobem poškozený. Operátor používá pro mapování tónů křivku zohledňující reálnou adaptaci zraku na světlo a uvažuje i realistickou maladaptaci[4] zapřičiněnou poškozenou lidskou nervovou soustavou, například v důsledku ztráty fotocitlivého pigmentu v čípcích či tyčinkách. Výsledkem tedy mohou být sekvence, které odpovídají vnímaní scény zdravým člověkem, či sekvence, jež simulují určité zrakové postižení jedince.

[4] Maladaptace – nevhodné přizpůsobení.

- **Virtual Exposures TMO** [6] cílí na explicitní potlačení šumu v produkovaných videích a vyznačuje se použitím dvou na sobě nezávislých bilaterálních filtrů. První je aplikován na jednotlivé snímky staticky za účelem rozdělení vstupního obrazu na vrstvu osvětlení a vrstvu detailů. Druhý filtr pracuje v temporální doméně a slouží pro cílenou redukci obrazového šumu. Dynamický rozsah základové vrstvy osvětlení je komprimován pomocí logaritmické tónovací křivky inspirované Dragovým operátorem mapování tónů [21].

- **Display Adaptive TMO** [54] popisuje mapování tónů jako problém optimalizace podobnosti vstupního HDR obrazu a výstupního videa zobrazeného na specifickém displeji. Hlavním cílem je minimalizovat rozdíl mezi kontrastem původní scény a výstupním kontrastem po provedení mapování tónů. Tónovací křivka používaná v tomto algoritmu mapuje vstupní hodnoty jasů do rozsahu, který lze využít pro stanovený model specifického displeje.

- **Temporal Coherence TMO** [7] řeší kromě jasových artefaktů také temporální problémy spojené s kontrastem. Po provedení mapování tónů proběhne dodatečné porovnání výsledků s původními daty. Autoři tvrdí, že není přípustné, aby temporální vlastnosti vstupních snímků nebyly ve výsledném videu zachovány. Tento požadavek lze ilustrovat jednoduchým příkladem. Operátor nepřipouští, aby některý ze snímků celé výsledné sekvence disponoval vyšší střední hodnotou jasu oproti snímku, který má ve vstupní sekvenci tuto hodnotu nejvyšší. Uplatnění tohoto principu sice umožňuje zachování reálných vlastností světla, nicméně výstupem operátoru jsou často snímky s nízkými hodnotami jasu, a výsledné sekvence pak mají velmi tmavý charakter.

- **Zonal Temporal Coherence TMO** [9] rozšiřuje operátor popsaný v předchozím bodě, přičemž provádí rozdělení obrazu do „zón“ na základě celkového histogramu jasu v logaritmické doméně. Nejprve algoritmus stanoví lokální maxima podle histogramu a lokální minima následně použije jako hranici pro oddělení jednotlivých zón. Tyto klíčové hodnoty se ovšem mohou napříč sekvencí skokově měnit, a proto operátor používá konstantní podobu hranic pro všechny snímky. Z důvodu zamezení vzniku obrazových artefaktů v oblasti přechodů mezi zónami jsou hodnoty v těchto oblastech průměrovány pomocí škálovacího koeficientu.

- **Motion Path Filtering TMO** [4] klade zvláštní důraz na maximální reprodukci lokálního kontrastu při zachování temporální koherence. Základním stavebním kamenem tohoto operátoru je odhad pohybu pro jednotlivé pixely počítaný pomocí optického toku. Další zajímavostí je výpočet základové vrstvy obrazu prostřednictvím filtrace v prostorově temporální doméně, zatímco detaily nesoucí vrstva vzniká pouze pomocí temporálního filtrování. Pozitivním vedlejším účinkem prostorově temporálního filtrování je mimo jiné i mírná redukce šumu ve výsledném videu.

- **Hybrid TMO** [73] předkládá unikátní metodu, kdy pro každý snímek ze sekvence používá dvě různé tónovací křivky. Obraz je nejprve rozdělen do dvou typů oblastí. První z nich tvoří části s velkou lokální změnou jasových hodnot a druhou oblasti jasově konstantní. Pro první typ oblastí následuje lokální mapování tónů pomocí operátoru iCAM06 [44]. Zbylé části obrazu jsou modifikovány podle podoby histogramu snímku. Aby bylo zamezeno vzniku temporálních artefaktů, probíhá pro každý snímek škálování jasu pomocí prahové funkce.

- **Noise-aware TMO** [25] cílí na produkci temporálně koherentního videa, které i při vysokém kompresním poměru dynamického rozsahu dokáže zachovat co nejvíce možných detailů bez projevu šumu způsobeného snímačem. Při výpočtu tónovací křivky jsou zohledněny lokální atributy scény, úrovně šumu a displej, na kterém je výstup zobrazován. Operátor také používá vlastní metodu pro filtrovaní zachovávající hrany k rozdělení obrazu na základovou a detaily nesoucí vrstvu, která zaručuje redukci artefaktů vznikajících převrácením gradientu kolem hran objektů ve scéně. Velkou výhodou je také možnost zpracování vstupního obrazu v reálném čase.

Kromě výše uvedených operátorů existuje celá řada dalších zajímavých algoritmů pro mapování tónů videa. Jako základní operátor pro implementační část této diplomové práce byl vybrán inovativní operátor *ELTM* publikovaný v roce 2018, který je velmi podrobně dekomponován a popsán v nadcházející části textu.

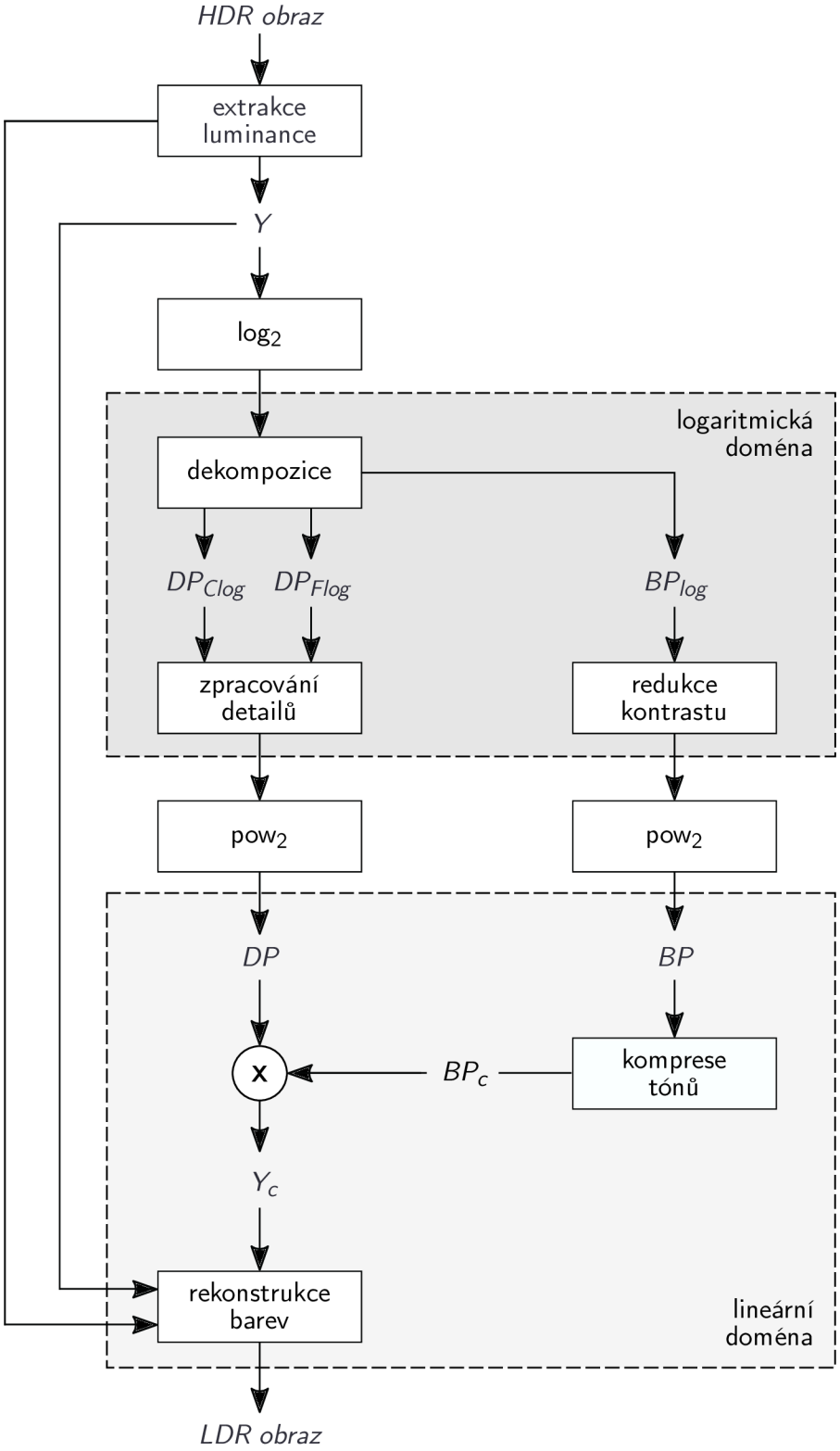
4.5 Operátor ELTM

Enhanced Local Tone-mapping Operator [27] (zkráceně *ELTM*) zesiluje lokální detaily obrazu a využívá maximální možný výstupní dynamický rozsah, díky čemuž produkuje výborný globální kontrast ve výsledných snímcích. Formálně se tento algoritmus řadí mezi lokální operátory, nicméně během výpočtu dochází k mnoha operacím i na globální úrovni. Kombinací lokálního a globálního přístupu je možné dosáhnout velmi zajímavých výsledků. Podobu výstupního videa lze ovlivnit prostřednictvím řady uživatelských parametrů, jež vytváří velmi intuitivní rozhraní.

4.5.1 Obecný popis algoritmu ELTM

Operátor ELTM slouží jako základ pro implementaci programu, jenž je součástí této diplomové práce, a je vhodné jej detailně rozebrat. Nejprve se tento text věnuje dílčím krokům mapování tónů jednotlivých snímků, na které dále navazuje podrobný popis temporálních operací.

Stanovená vrstva relativní luminance Y je převedena do logaritmické domény o základu 2 a následuje její dekompozice na základovou (BP_{log}) a dvě detaily nesoucí vrstvy (DP_{Clog}, DP_{Flog}) pomocí řízeného filtru. Dynamický rozsah základové vrstvy je poté modifikován lineárním škálováním, které posune rozsah do předem definovaného intervalu. Vrstvy detailů jsou zpracovány nezávisle na sobě. DP_{Clog} představuje rovinu „hrubších“ detailů (angl. *coarse details*), zatímco vrstva DP_{Flog} nese detaily „jemnější“ (angl. *fine details*). Toto rozdělení umožňuje separátní kontrolu zesílení hrubých a jemných detailů. Následuje komprese tónů v lineární doméně pomocí adaptivní tónovací křivky, která získává svoji podobu na základě statistických údajů snímku. Tato komprese je provedena pouze pro základovou vrstvu BP_{log}, jež je následně vynásobena oběma vrstvami detailů. Vzniká tak komprimovaná jasová vrstva Y_c, která odpovídá výslednému dynamickému rozsahu. V posledním kroku probíhá obnovení barev jednotlivých kanálů. Blokové schéma operátoru ELTM pracujícího staticky pro jeden snímek je uvedeno na obrázku 4.5.



Obrázek 4.5: Blokové schéma operátoru ELTM.

4.5.2 Stanovení relativní luminance

Prvním krokem pomyslného řetězce zpracování je výpočet relativní luminance konkrétního snímku. Přestože některé operátory mapují tóny každého barevného kanálu zvlášť, algoritmus ELTM používá tradiční způsob, kdy komprese rozsahu probíhá pouze v jasové doméně, která obsahuje bezbarvé neboli achromatické informace o jasových hodnotách snímku. Tato jasová vrstva je komprimována a upravena tak, že je z ní možné v závěrečné fázi algoritmu obnovit původní barevnost snímku. Relativní luminanci lze vypočítat z hodnot jednotlivých barevných kanálů obrazu pomocí rovnice, kterou definuje sada doporučení ITU BT. 709 (viz 3.2.1):

$$Y = 0.2126 \cdot R + 0.7152 \cdot G + 0.0722 \cdot B. \tag{4.12}$$

Rovnici 4.12 lze nezávisle aplikovat na všechny obrazové body vstupního snímků, čímž vznikne vrstva, na kterou se tento text dále odkazuje jako na vrstvu relativní luminance Y. Proměnné R, G, B vyjadřují červený, zelený a modrý barevný kanál.

4.5.3 Dekompozice obrazu

Prvním krokem dekompozice je převod jasové vrstvy do logaritmické domény, kde hodnoty obrazových bodu lineárně korelují s množstvím světla, které bylo snímacím čipem fotoaparátu či kamery zachyceno. Výhodou využití logaritmické domény je skutečnost, že jasové hodnoty zde mohou být rozděleny podle frekvence, kde nízkofrekvenční hodnoty vyjadřují celkové osvětlení scény, zatímco vysokofrekvenční hodnoty představují reflektanci (viz 4.1) odpovídající odrazivosti objektů, a tedy detailům obrazu [23]. Převod jednotlivých pixelů snímku do logaritmické domény o základu 2 lze provést podle podle vzorce 4.13:

$$Y_{log} = \log_2(Y + \varepsilon), \tag{4.13}$$

kde Y_{log} představuje jasovou vrstvu v logaritmické doméně a symbol ε představuje konstantu o hodnotě 10^{-6}. Tato konstanta zabraňuje problémům při použití logaritmu na nulové hodnoty, pro které není logaritmus definovaný.

Pro další krok dekompozice jasové vrstvy v logaritmické doméně je nutné použít některý z filtrů, který zachovává hrany. Operátor ELTM pro tento účel používá řízený filtr popsaný v sekci 4.1.3. I přestože zde lze zvolit některý ze sofistikovanějších filtrů, řízený filtr přináší velmi kvalitní výsledky a poskytuje příznivou výpočetní složitost. Nabízí se zde i možnost využití bilaterálního filtru, nicméně jeho složitost a náchylnost k artefaktům způsobených převrácením gradientu ukazují, že řízený filtr je vhodnější variantou.

Namísto klasického přístupu, kdy je obraz dekomponován do jedné základové a jedné detaily nesoucí vrstvy, je zde zvolen mírně odlišný přístup využívající dvou rovin detailů. První z těchto vrstev, značená jako DP_{Clog}, reprezentuje detaily hrubého měřítka a lze ji vypočítat pomocí filtru s prostorově větším konvolučním jádrem. Tato rovina obsahuje detaily spíše rozsáhlejších rozměrů, které vizuálně odpovídají lokální odrazivosti a kontrastu celých objektů ve scéně. Druhá vrstva DP_{Flog} naproti tomu vzniká filtrací pomocí malého konvolučního jádra a představuje reflektanci hran objektů. Nese také vizuální data reprezentující drobné nerovnosti povrchu objektů.

Toto rozdělení na dvě úrovně detailů umožňuje operátoru ELTM velmi robustní kontrolu nad zesílením lokálního kontrastu. Kupříkladu vzájemný kontrast jednotlivých objektů ve scéně lze výrazně zesílit, zatímco jemné detaily mohou být zesíleny pouze lehce, aby nedocházelo k nepřiměřené ostrosti výsledného obrazu, či k amplifikaci nežádoucího šumu.

Během dekompozice je možné omezit amplitudu hodnot vrstev detailů pomocí parametru λ_C pro hrubé detaily a parametru λ_F pro detaily jemné. Parametry λ_C, λ_F definují, jaké množství detailů bude extrahováno z původní vrstvy logaritmické relativní luminance Y_{log}. Motivace pro použití těchto omezujících parametrů tkví v možnosti omezení skokových změn v původním obraze způsobených například impulzním šumem či odrazem světla od lesklých povrchů objektů. Tento přístup přináší velkou flexibilitu algoritmu a lze pomocí něj operátor konfigurovat adekvátně vůči konkrétní scéně. Při nastavení parametrů λ_C, λ_F na nulové hodnoty lze potlačit extrakci detailů, a degradovat tak operátor ELTM na globálně fungující algoritmus.

Při dekompozici nejprve probíhá postupná extrakce detailů z vrstvy Y_{log}. V první fázi jsou vyňaty jemné detaily a následně detaily hrubšího měřítka, čímž vzniká základová vrstva obrazu. Proces dekompozice na zmíněné obrazové vrstvy lze matematicky vyjádřit prostřednictvím následující série operací 4.14:

$$\begin{aligned} DP_{Flog} &= \mathrm{clip}(Y_{log} - \mathrm{GF}(Y_{log}, r_s, \varepsilon_s), \lambda_F), \\ BP_{Flog} &= Y_{log} - DP_{Flog}, \\ DP_{Clog} &= \mathrm{clip}(BP_{Flog} - \mathrm{GF}(BP_{Flog}, r_l, \varepsilon_l), \lambda_C), \\ BP_{log} &= BP_{Flog} - DP_{Clog}. \end{aligned} \tag{4.14}$$

Hodnoty r_s, r_l představují poloměr malého (r_s), resp. velkého (r_l) konvolučního jádra a jsou použity jako parametry pro řízený filtr GF. Regularizační parametry filtru reprezentují konstanty $\varepsilon_s, \varepsilon_l$ (viz rovnici 4.6). BP_{Flog} a BP_{log} vyjadřují základovou vrstvu osvětlení po odečtení jemných, resp. jemných i hrubých, detailů od logaritmické vrstvy relativní luminance. Funkce $\mathrm{clip}(range, limit)$ slouží pro odstranění (přesněji vynucenou saturaci) hodnot z rozsahu $range$ překračující práh daný rozsahem $[-limit, limit]$.

V rámci algoritmu ELTM byla pro poloměr malého konvolučního jádra r_s zvolena konstantní velikost 3. Velikost poloměru většího jádra r_l pro filtraci hrubých detailů je stanovena pomocí vzorce 4.15:

$$r_l = 0{,}1 \cdot \min(width, height), \tag{4.15}$$

kde $width$ reprezentuje šířku a $height$ výšku vstupního snímku, a výsledek rovnice odpovídá jedné desetině kratšího rozměru vstupního snímku. Oba dva regularizační parametry řízeného filtru, $\varepsilon_s, \varepsilon_l$, byly stanoveny na konstantní hodnotu 0,1. Limit amplitudy hrubých detailů λ_C je nastaven na vysokou výchozí hodnotu 1, kdy nedochází k žádné omezující limitaci. Naproti tomu výchozí hodnota pro limit jemných detailů λ_F je stanovena na 0,02, kdy lze efektivně potlačit impulzní šum a intenzivní odlesky od povrchů s vysokou mírou odrazivosti světla.

4.5.4 Komprese rozsahu kontrastu základové vrstvy

Častým přístupem pro kompresi kontrastu je jednoduché škálování základové vrstvy v logaritmické doméně s použitím konstantního škálovacího činitele [23, 79], což je ekvivalentní operace vůči aplikaci mocninné funkce v lineární doméně. Je-li komprese tónů provedena čistě pomocí škálování v logaritmické doméně, mohou větší kompresní poměry vést ke ztrátě kontrastních informací.

Další technikou může být posun logaritmických jasových hodnot do předem stanoveného rozsahu za účelem následné redukce kontrastu v lineární doméně. Tento přístup sice zachovává dobrý poměr globálního kontrastu, ale často vede ke ztrátě detailů ve světlých oblastech snímku [64].

Aby bylo možné co nejlépe reprodukovat globální kontrast a aby zároveň nedocházelo ke ztrátě detailů ve světlých oblastech snímku, kombinuje operátor ELTM obě dvě techniky zmíněné v předchozích odstavcích. Komprese tónů je prováděna částečně v logaritmické i lineární doméně. Redukce kontrastu v logaritmické doméně slouží pro převod hodnot do předem stanoveného rozsahu, zatímco následná komprese tónů v lineární doméně přizpůsobuje kontrast podle určitých parametrů operátoru a zodpovídá za výsledný vzhled snímku. Tento přístup je důležitý zejména z důvodu zvýšení celkové robustnosti algoritmu a konzistentního chování operátoru pro širokou škálu vstupních dat.

Jelikož rozsah vstupního obrazu se může pro každý snímek lišit velmi razantně, hrozí, že nemusí dojít k efektivnímu pokrytí celého potenciálu výstupního dynamického rozsahu. Kvůli této skutečnosti operátor ELTM provádí kompresi rozsahu do předem definovaného intervalu $[-\tau_R, 0]$:

$$\beta = -\max(BP_{log}), \quad \alpha = \frac{\tau_R}{\max(BP_{log}) -, \min(BP_{log})}, \tag{4.16}$$

$$BP'_{log} = \alpha \cdot (BP_{log} + \beta), \tag{4.17}$$

kde α představuje škálovací činitel a β vyjadřuje kompenzační činitel sloužící pro posun hodnot v logaritmické doméně. BP'_{log} označuje základovou vrstvu po provedení škálování a posunutí hodnot. Cílem těchto operací je umístění nejvyšších logaritmických hodnot jasu na hodnotu 0, díky čemuž lze kompenzovat nevhodnou expozici vstupního snímku. Tato jednoduchá úprava zaručuje, že histogram logaritmické základové vrstvy BP'_{log} odpovídá předdefinovanému rozsahu hodnot a je možné následně provést kompresi v lineární doméně. Jako cílový rozsah byl zvolen interval [-5, 0]. Pro další zvýšení robustnosti operátoru se namísto maximální a minimální hodnoty, reprezentované funkcemi max a min v rovnicích 4.16, používají hodnoty získané pomocí percentilu o velikosti 99,9, resp. 0,01.

4.5.5 Zpracování detailů

Vrstvy obsahující detaily jsou upraveny v logaritmické doméně pomocí uživatelem definovaných parametrů pro zesílení či potlačení jejich hodnot. Na zesilování detailů existují dva různé pohledy. První z nich pohlíží na amplifikaci spíše skepticky, protože mohou být velmi nepříznivě ovlivněny šumem [6], zejména v tmavých oblastech snímků. Druhý přístup naopak nabádá k zesilování detailů, a to zejména proto, že běžné displeje disponují velmi omezenými zobrazovacími schopnostmi pro nízké jasové hodnoty [54].

Operátor ELTM předpokládá, že problémy související se zpracováním šumu jsou adresovány mimo jeho algoritmus, a používá druhý ze zmíněných přístupů. Pro posílení detailů v tmavých oblastech využívá prostorově proměnné zpracování, jehož kontrolu značně usnadňuje skutečnost, že v této fázi se hodnoty pixelů nacházejí v předem známém intervalu. Jestliže intenzita jasu jednotlivých pixelů nabývá hodnot z rozsahu [-5, 0], kde jsou tmavé oblasti reprezentovány nízkými hodnotami, lze vypočítat proměnlivou masku zesílení vynásobením základové vrstvy konstantou o velikosti -0,4. Po provedení této operace vzniká vrstva SG, která nabývá hodnot od 0 pro nejsvětlejší oblast po 2 pro oblast nejtmavší, protože při násobení rozsahu konstantou o záporné hodnotě dochází zároveň k jeho inverzi. Nežádoucí zesílení detailů pro vysoké hodnoty jasů lze eliminovat pomocí odstranění hodnot masky o velikosti menší než 1:

$$SG = \max(-0{,}4 \cdot BP'_{log}, 1), \tag{4.18}$$

$$DP'_{Clog} = \eta_C \cdot SG \cdot DP_{Clog}, \quad DP'_{Flog} = \eta_F \cdot SG \cdot DP_{Flog}. \tag{4.19}$$

Konstanty η_C, η_F z rovnice 4.19 představují prostorově uniformní zesilovací činitele sloužící pro amplifikaci hrubých a jemných detailů. DP'_{Clog} a DP'_{Flog} vyjadřují roviny hrubých a jemných detailů po provedení zesílení. Ve výchozí konfiguraci je uživatelsky volitelný parametr pro amplifikaci hrubých detailů η_C nastaven na hodnotu 1,5 z důvodu posílení lokálního kontrastu, zatímco parametr zesilující jemné detaily η_F je nastaven na hodnotu 1. Tato kombinace zajišťuje pro většinu vstupních sekvencí velmi dobrý poměr mezi intenzitou lokálního kontrastu a projevem šumu na výsledných snímcích.

Po aplikaci zmíněných operací je nutné provést převod vrstev zpět do lineární domény pomocí mocninné funkce o základu 2:

$$BP = 2^{BP'_{log}}, \quad DP = 2^{(DP'_{Clog}+DP'_{Clog})}. \tag{4.20}$$

Proměnná BP z rovnic 4.20 představuje základovou vrstvu v lineární doméně. Finální lineární vrstvu DP kombinující obě roviny detailů lze získat jejich sečtením a následnou aplikací mocninné funkce. V dalším kroku algoritmu probíhá komprese tónů linearizované základové vrstvy.

4.5.6 Komprese tónů

Po převedení základové a detaily nesoucí vrstvy do lineární domény následuje jako další krok operátoru ELTM komprese tonality. Díky úpravě rozsahu hodnot základové vrstvy v logaritmické doméně je tento krok méně závislý na specifickém rozsahu jednotlivých snímků. Samotnou redukci rozsahu a zesílení kontrastu lze provést pomocí dobře známých globálních funkcí, jako jsou například gamma, nebo logaritmická, či racionální kompresní funkce [82]. Weberův-Fencherův zákon říká, že intenzita zrakového vjemu je logaritmicky závislá na intenzitě stimulujícího podnětu [17]. Z této skutečnosti vychází i premisa, že kompresi základové vrstvy lze provést pomocí vzorce 4.21:

$$BP_c = \frac{\log(BP + p) - \log(p)}{\log(1 + p) - \log(p)}, \tag{4.21}$$

kde BP_c představuje komprimovanou základovou vrstvu a p reprezentuje parametr ovlivňující sílu komprese. Při použití této jednoduché operace ovšem hrozí ztráta velkého množství globálního kontrastu, kdy výsledné snímky mohou působit „mdle“. Za účelem lepšího zachování celkového kontrastu lze využít škálování hodnot do výstupního rozsahu, nicméně pokud je aplikováno až po provedení komprese tónů, dojde ke ztrátě informací, protože tónovací křivka nerespektuje celkový potenciál výsledného rozsahu.

Aby bylo možné využít tónovací křivku co nejefektivněji a nedocházelo k vypouštění informací, operátor ELTM aplikuje kompresi tónů na základě statistik získaných ze základové vrstvy. Tento postup zajišťuje velmi dobrý globální kontrast a zároveň umožňuje využít celý výstupní dynamický rozsah. Při použití vzorce 4.21 by mohlo dojít k nechtěnému potlačení detailů ve velmi světlých i tmavých oblastech snímku. Za účelem zachování co největšího možného lokálního kontrastu i v těchto částech obrazu je navrženo omezení výstupního dynamického rozsahu, aby nedocházelo k potlačování detailů ve zmíněných citlivých oblastech. Modifikovanou rovnici zahrnující omezení výstupního rozsahu i statistické parametry získané z obrazu lze vyjádřit následujícím způsobem:

$$BP_c = (BP_{cmax} - BP_{cmin}) \cdot \frac{\log(\frac{BP - BP_{min}}{BP_{max} - BP_{min}} + p) - \log(p)}{\log(1 + p) - \log(p)} + BP_{cmin}. \tag{4.22}$$

Minimální a maximální hodnoty základové vrstvy, reprezentovány proměnnými BP_{min} a BP_{max}, jsou vypočítány pomocí percentilů o velikostech 0,1 a 99,9. Uživatelem nastavitelné parametry BP_{cmin} a BP_{cmax} představují minimální a maximální hodnoty výstupního rozsahu. Jejich výchozí hodnoty byly stanoveny na 0,1 a 0,9.

Samotná komprese tónů není omezena pouze na výpočet pomocí rovnic 4.21 či 4.22. Lze je nahradit celou řadou jiných globálních metod. Výše uvedené postupy ovšem umožňují velmi přímočarou kontrolu celkového jasu výsledného snímku a poskytují velmi kvalitní výsledky. Navíc nevyžadují žádné zásahy do základové vrstvy, vyjma škálování rozsahu hodnot do předem stanoveného intervalu. Některé operátory mapování tónů ve svých algoritmech využívají kompresi v logaritmické doméně [23, 54], některé v lineární [69, 79]. Například velmi populární operátor pro statické snímky s názvem *Photographic Tone-mapping Operator* [64] používá stejně jako ELTM kombinaci těchto přístupů a také poskytuje výborný výsledný obraz.

4.5.7 Obnovení detailů a barev

Po provedení operací nad základovou vrstvou je dále nutné obnovit detaily obrazu, což lze provést pomocí jednoduchého vynásobení komprimované základové vrstvy BP_c detaily nesoucí vrstvou DP. Vzniká tak vrstva Y_c, která obsahuje hodnoty jasu celého snímku po aplikaci mapování tónů:

$$Y_c = BP_c \cdot DP. \tag{4.23}$$

V posledním kroku algoritmu ELTM probíhá obnovení barev z lineárních hodnot jednotlivých barevných kanálů původního HDR obrazu. Výpočet probíhá na základě Schlickova vzorce [69], jenž je doplněn o uživatelem nastavitelný parametr s, umožňující modifikovat barevnou saturaci výsledného obrazu [79]:

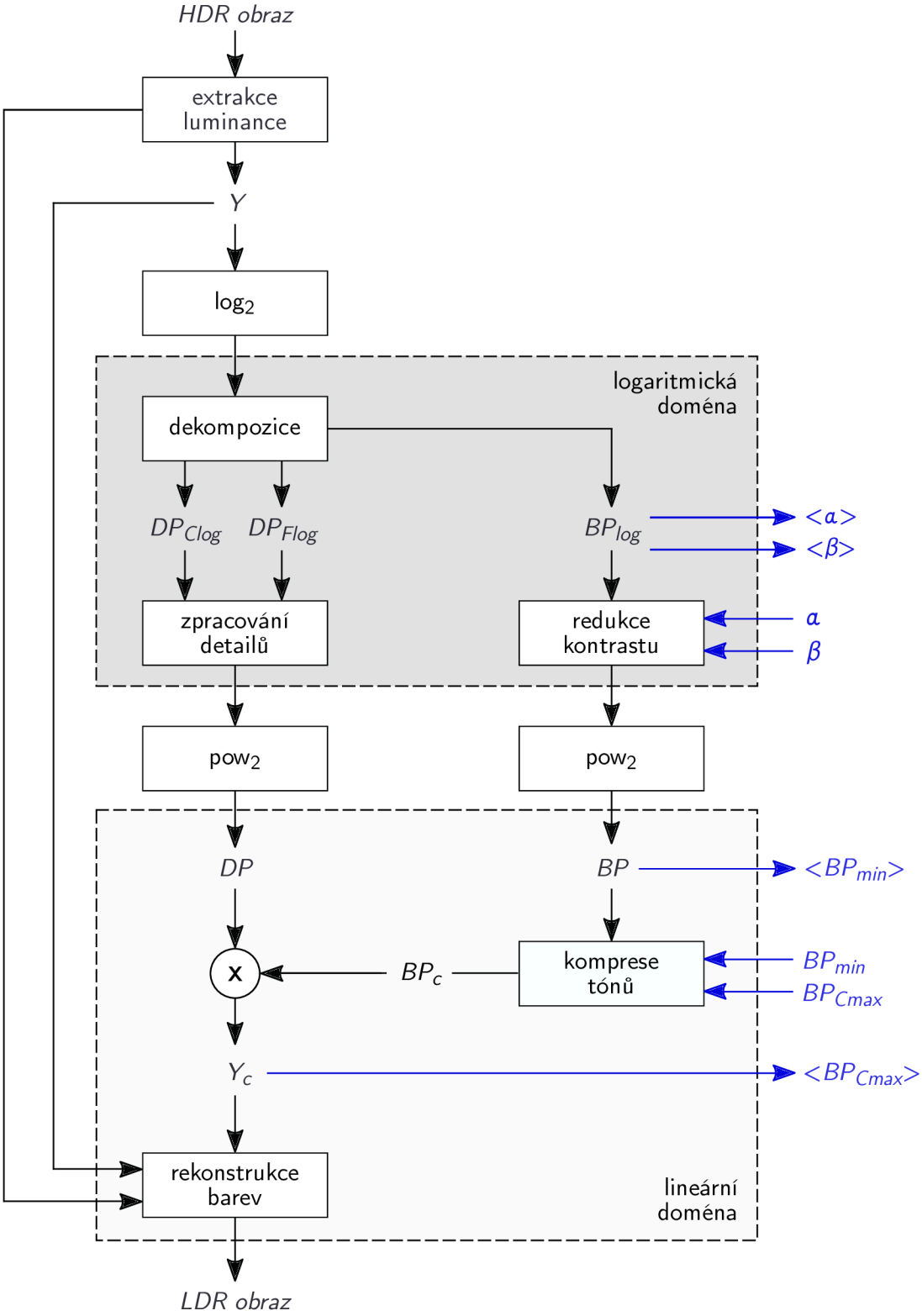
$$C_{out} = Y_C(\frac{C_{in}}{Y_{in}})^{\frac{s}{\gamma}},\ C \in R, G, B. \tag{4.24}$$

Parametr γ z rovnice 4.24 představuje činitel o konstantní hodnotě 2,2, jenž je nutný pro korektní výpočet hodnot odpovídajících barevnému prostoru sRGB. Tato exponenciální operace se obecně nazývá *gamma korekce* a slouží pro kalibraci obrazových dat za účelem jejich zobrazení na displeji [83]. Pomocí proměnné s může uživatel stanovit míru barevné saturace a její výchozí velikost je nastavena na neutrální hodnotu 1.

4.5.8 Temporální zpracování snímků

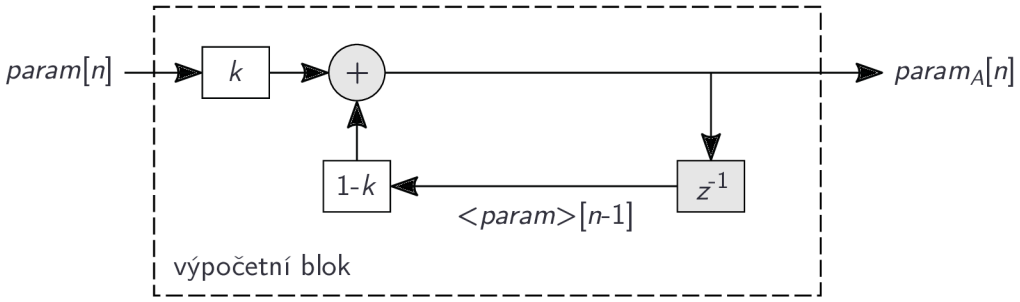
Jelikož operátor ELTM poskytuje pro statické obrázky velmi dobré výsledky v relativně krátkém čase výpočtu, rozhodli se autoři rozšířit jeho funkcionalitu tak, aby bylo možné jej použít i pro videosekvence [28]. Cílem bylo navrhnout univerzální řešení zajišťující dobrý poměr mezi temporální koherencí a temporální adaptací (viz 4.3.2) ve výsledném videu.

Jednotlivé parametry, které operátor ELTM používá v rámci svého algoritmu, lze pomyslně rozdělit do dvou kategorií. V prvním případě lze hovořit o uživatelsky konfigurovatelných proměnných, které zůstávají konstantní pro všechny snímky ze vstupní sekvence. Druhou kategorií jsou statistické parametry, jejichž výpočet probíhá pro každý snímek zvlášť na základě aktuálního snímku a statistik získaných ze snímků předchozích. Nyní je vhodné uvést podrobné schéma algoritmu ELTM obohacené o temporální parametry:



Obrázek 4.6: Schéma operátoru ELTM doplněné o temporální parametry. Parametry ve špičatých závorkách reprezentují statistické hodnoty použité pro předchozí snímek ze vstupní sekvence, které operátor zohledňuje při výpočtu jejich nových hodnot pro aktuální snímek.

Stanovení temporálně adaptovaných hodnot statistických parametrů pro aktuální snímek probíhá v samostatném výpočetním bloku. Jeho vstupem je konkrétní hodnota daného parametru získaná ze současného snímku. Druhou proměnnou, kterou blok uvažuje, je hodnota parametru, která byla použita pro předchozí snímek, značená pomocí špičatých závorek ve tvaru <*param*>. Výstupem bloku je hodnota parametru, kterou algoritmus ELTM dále použije pro zpracování aktuálního snímku. Funkcionalita výpočetního bloku odpovídá jednoduchému diskrétnímu filtru s nekonečnou impulzní odezvou typu IIR[5] prvního řádu, jenž zohledňuje při svém výpočtu kumulativní statistickou hodnotu získanou z předchozích snímků:



Obrázek 4.7: Schéma popisující vztah operátoru mapování tónů a výpočetního bloku představujícího filtr typu IIR, který zodpovídá za výpočet statistických parametrů pro aktuálně zpracovávaný snímek.

Hodnota *param* vyjadřuje hodnotu obecného parametru, která je vypočtena pro aktuální snímek s pořadovým číslem n před aplikací temporální adaptace. Ta je vynásobena koeficientem k, jenž je pro daný parametr v rámci celé sekvence konstantní. Dalším krokem je přičtení hodnoty parametru, jež byla použita pro předchozí snímek, vynásobenou koeficientem o velikosti $1 - k$. Výstupem filtru, jehož schéma je uvedeno na obrázku 4.7, je pak hodnota $param_A$, jež vyjadřuje temporálně adaptovanou hodnotu parametru, kterou algoritmus reálně použije v dalších krocích výpočtu. Matematický zápis pro výpočet temporálně adaptované hodnoty obecného statistického parametru lze vyjádřit pomocí následující rovnice 4.25:

$$param_A[n] = k \cdot param[n] + (1 - k) \cdot param_A[k - 1]. \tag{4.25}$$

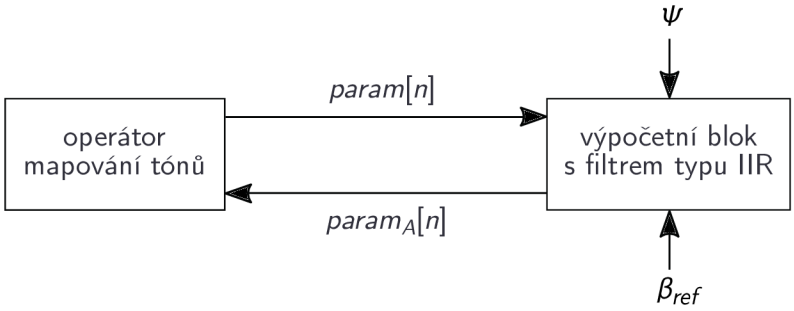
Dodatečnou regulaci míry temporální koherence kontrastu lze provést pomocí parametru β, jenž odpovídá za posun dynamického rozsahu před jeho samotnou kompresí. Za účelem zachování vyšší míry koherence výstupního videa může uživatel vybrat jeden snímek ze sekvence a označit jej jako referenční. Vybraný snímek slouží pro výpočet výchozí hodnoty parametru β, označené jako β_{ref}, kterou algoritmus bere v úvahu při výpočtu konkrétní hodnoty β_{Aref} pro každý snímek. Nová hodnota β_{Aref} je pro jednotlivé snímky vypočtena podle rovnice 4.26:

$$\beta_{Aref} = \Psi\beta_A + (1 - \Psi)\beta_{ref}, \tag{4.26}$$

kde β_{ref} reprezentuje výchozí hodnotu statistického parametru β získanou z referenčního snímku pomocí vztahu 4.16. Proměnná Ψ představuje uživatelský parametr, který ovlivňuje dodatečnou míru koherence již adaptovaného parametru β_A vůči referenční hodnotě β_{ref}. Rozsah, ze kterého lze zvolit parametr Ψ, je tvořen intervalem $[0, 1]$. Hodnota 0 znamená

[5]Filtr typu IIR: filtr s nekonečnou impulzní odezvou (angl. *infinite impulse response*).

maximální koherenci, protože pro každý snímek ze sekvence bude použita stejná hodnota $\beta_{Aref} = \beta_{ref}$, zatímco hodnota 1 dovoluje zachovat maximální zvolenou adaptaci statistického parametru β pro každý snímek. Při požadavku na zpracování videa v reálném čase lze jako referenční snímek nastavit hned první snímek ze vstupní sekvence. Kompletní schéma výpočetního bloku, jež zohledňuje při výpočtu zvolený referenční snímek, lze zakreslit následovně:



Obrázek 4.8: Schéma rozhraní výpočetního bloku pro temporální výpočty.

Výpočetní blok si uchovává v paměti hodnoty jednotlivých statistických parametrů použitých operátorem ELTM při zpracování předchozího snímku s pořadovým číslem $n-1$. Parametry Ψ a β_{ref} mají vliv pouze na výpočet nové hodnoty statistiky β (resp. β_{Aref}) a při výpočtu hodnot ostatních parametrů nejsou brány v úvahu.

Výpočet statistické hodnoty BP_{cmax} probíhá v rámci algoritmu interně a uživatel nemusí její velikost definovat explicitně. Tato proměnná umožňuje zachování detailů v nejsvětlejších oblastech scény. Jestliže výsledná vrstva jasu (Y_c) obsahuje hodnoty menší než 1, nedochází k využití celého potenciálu výstupního dynamického rozsahu. Pro následující snímek, který se vyskytuje v sekvenci vyskytuje po aktuálně zpracovávaném snímku, lze stanovit novou hodnotu statistiky BP_{cmax} podle rovnice:

$$BP_{cmax} = \max(\frac{1}{Y_{cmax}},\ 1). \tag{4.27}$$

Zvýšení horní hranice rozsahu, ze kterého je počítána komprimovaná základová vrstva osvětlení BP_c (viz rov. 4.22), způsobí, že nedojde k vypuštění detailů v oblastech s vysokými hodnotami jasu. Funkce max z předchozí rovnice 4.27 vybere vždy vyšší hodnotu z dvojice vstupních proměnných. První proměnnou představuje inverzní hodnota maximálního jasu vrstvy Y_c. Druhou proměnnou je konstanta o velikosti 1. Aby rapidní změny maximální hodnoty jasu nezpůsobovaly jasové artefakty, probíhá následná kompenzace aplikované hodnoty BP_{cmax} pomocí filtru typu IIR prvního řádu, stejně jako je tomu v případě ostatních statistických hodnot.

Kapitola 5

Návrh a implementace aplikace

Nedílnou součástí teto diplomové práce je kromě teoretické části také implementace vybraného operátoru mapování tónů. Aplikace musí s použitím zvoleného algoritmu umožňovat mapování tonality v temporální doméně a poskytovat kvalitní výsledky bez projevu obrazových artefaktů. Po konzultaci s vedoucím práce a důkladné analýze současného stavu na poli moderních operátorů byl pro implementaci vybrán *Enhanced Local Tone-mapping Operator* [27], jenž byl detailně rozebrán v sekci 4.5. Vzhledem k tomu, že se jedná o netriviální temporální operátor publikovaný v roce 2018, na internetu ještě pořád neexistuje jeho dostupná implementace (ke květnu roku 2019). Dalším argumentem, jenž hovoří pro zvolení této metody, je její příznivá výpočetní náročnost a možnost použití v reálném čase. Unikátnost operátoru ELTM je také dána použitím řízeného filtru pro extrakci základové vrstvy osvětlení scény a dvou na sobě nezávislých rovin detailů. Nad rámec zadání bude vypracováno přehledné grafické rozhraní, které uživateli poskytne okamžitou zpětnou vazbu při nastavování jednotlivých parametrů pro samotné temporální mapování tónů.

5.1 Požadavky na funkcionalitu aplikace

Program musí být schopen načíst vstupní sérii snímků reprezentující video s vysokým dynamickým rozsahem. Podporované vstupní formáty snímků budou OpenEXR (*.exr) a Radiance RGBE (*.hdr), které byly popsány v sekci 3.1. Po provedení konfigurace uživatelských parametrů operátoru bude aplikováno temporální mapování tónů, čímž vznikne výstupní video. Pro výstupní video byl vybrán kontejner MP4 a formát H.264, protože tato kombinace produkuje nejpoužívanější a nejpřenositelnější video soubory s běžným dynamickým rozsahem. Bylo by sice možné využít některý z modernějších formátů, který by byl schopen zajistit menší datovou velikost výsledného videa, nicméně by pro tento účel bylo nutné v rámci implementace použít specializovanou knihovnu pro kódování a přenositelnost videa by byla výrazně horší.

Aplikace bude schopna provést mapování tónů na libovolnou sekvenci snímků, kde data odpovídají specifikaci zmíněných vstupních formátů. Protože podoba dat může být pro různé sekvence velmi odlišná, je nutné uvážit i jejich předzpracování pro následné mapování tonality. Požadavky na samotné mapování tónů vychází z teoretické části textu. Je zejména potřeba zajistit, aby výsledná videa neobsahovala obrazové artefakty. Největší důraz bude kladen na temporální koherenci jasu vůči reprezentované scéně, potlačení problikávání a zamezení výskytu duchů.

5.2 Požadavky na uživatelské rozhraní

Přestože původní zadání nepočítá s návrhem grafického uživatelského rozhraní, bude v rámci této práce implementováno. Je důležité, aby uživatel měl možnost zkoumat vliv jednotlivých parametrů operátoru mapování tónů na výsledný obraz. Důraz bude kladen zejména na intuitivnost a co největší uživatelskou přívětivost, aby i uživatel, u kterého nelze předpokládat předchozí zkušenost s temporálním mapováním tonality, dokázal během krátké chvíle pochopit význam jednotlivých konfigurovatelných parametrů. Aplikace bude schopna zobrazit náhled zvoleného snímku po provedení mapování tónů. Operátor ELTM počítá se skutečností, že jeden snímek ze sekvence je vybrán jako referenční. Uživatel bude mít možnost intuitivně zvolit, který snímek bude k tomuto účelu zvolen. Aby byla zachována simplicita rozhraní, nemusí nutně zobrazené konfigurační prvky pro jednotlivé parametry přesně odpovídat jejich původním variantám, jež specifikuje teoretická část 4.5. Všechny modifikace rozhraní operátoru, které aplikace zahrnuje, jsou detailně popsány v rámci následující kapitoly věnující se vlastní implementaci.

5.3 Nástroje zvolené pro implementaci

Jako programovací jazyk byl vybrán objektově orientovaný jazyk C++ v jeho aktuálním standardu ISO C++17[1]. Pro návrh a implementaci grafického uživatelského prostředí byl zvolen aplikační rámec (neboli *framework*) Qt[2] ve verzi 5.12 LTS, který poskytuje kromě sady knihoven také vývojové prostředí Qt Creator ve variantě s otevřeným zdrojovým kódem. Jediná externí knihovna, která byla nad rámec knihoven jazyka C++ a frameworku Qt použita, se nazývá OpenCV 3[3]. Jedná se o knihovnu vydávanou se svobodnou licencí a je možné ji využít pro akademické i komerční účely. Nástroj OpenCV byl navržen pro rychlé výpočty matematických úloh používaných v oblasti zpracování obrazu a je implementován v programovacích jazycích C a C++.

Velkou výhodou této zvolené kombinace je přenositelnost programu na širokou škálu různých platforem. Všechny použité nástroje podporují běh vyvinutých aplikací na nejpoužívanějších operačních systémech, včetně linuxových systémů, Apple MacOS a Microsoft Windows. Výsledný program tedy lze jednoduše přenést mezi různými systémy, kdy po instalaci potřebných knihoven stačí pouze drobná úprava konfiguračního souboru projektu. Vývoj probíhal z největší části na noteboocích Apple MacBook Pro 2015 a 2017. Tyto stroje disponují velmi kvalitními displeji s výborným rozsahem kontrastu, které výrobce komerčně označuje jako „Retina“, což značně usnadnilo vizuální zkoumání obrazových vlastností výsledných videí.

Popis implementace lze rozdělit na dvě části. První úsek se věnuje předzpracování snímků a implementaci samotného temporálního mapování tónů. Ve druhé části je popsán vývoj grafického uživatelského rozhraní a jsou zde zdůrazněny rozdíly mezi rozhraním operátoru prezentovanou v původních článcích [27, 28] a variantou implementovanou v rámci této diplomové práce. Jelikož jsou rozdíly rozhraní operátoru popsány z pohledu uživatele, jsou zařazeny do úseku, jenž se věnuje uživatelskému prostředí.

[1] https://iso.org/standard/68564.html
[2] https://qt.io/
[3] https://opencv.org/releases/

5.4 Implementace funkcionality

Pro usnadnění práce s obrazovými daty slouží datový typ `cv::Mat` implementovaný knihovnou OpenCV. Tento datový typ podporuje základní operace nad maticemi pixelů, například jejich násobení jinými maticemi či skalárními hodnotami. Vzhledem k faktu, že knihovna OpenCV dokáže tyto operace provádět na procesoru paralelně, je tento přístup výrazně efektivnější oproti sekvenční aplikaci operací na každý bod matic zvlášť. Datový typ `cv::Mat` navíc umožňuje pracovat s různými typy matic. Pro účely implementace byly použity typy `CV_32FC3`, reprezentující matici 32bitových pixelů se třemi barevnými kanály, a `CV_32FC1` pro obrazy s jedním barevným kanálem a 32bitovou hloubkou jednotlivých obrazových bodů. Výstupní snímky jsou dočasně uloženy pomocí matice typu `CV_8UC3` nesoucí 3 barevné kanály s 8bitovou hloubkou, což odpovídá klasickému digitálnímu obrazu s nízkým (resp. běžným) dynamickým rozsahem.

Při popisu vlastní implementace je v první fázi rozebrána funkcionalita samotného operátoru mapování tónů, jelikož nejlépe navazuje na předchozí teoretickou část práce. Protože samotný operátor ELTM nepopisuje předzpracování obrazu, byla navržena temporální metoda, díky které je možné pro mapování tonality použít co největší možný dynamický rozsah vstupních obrazů. Předložená metoda také dokáže efektivně zamezovat globálním jasovým artefaktům a redukovat nežádoucí šum ve výsledném videu. Některé vstupní sekvence obsahují snímky se zápornými hodnotami pixelů, které způsobují problémy v oblasti vzájemné jasové koherence snímků. Jmenovitě jde o některé sekvence z testovací sady[4] dodané s článkem *Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays* [32]. Po vysvětlení metody pro předzpracování obrazu následuje popis implementace jednotlivých kroků statického mapování tonality a poslední část této podkapitoly se věnuje temporální adaptaci operátoru pro videosekvence.

Obecnou abstrakci pro mapování tónů videa představuje třída `ToneMapping`, jejíž rozhraní poskytuje možnost inicializovat konkrétní operátor a zapisovat výsledné snímky do výstupního video souboru. Tato třída byla navržena tak, aby bylo snadno možné případně rozšířit aplikaci o podporu dalších algoritmů pro mapování tonality. Funkcionalitu temporálního operátoru ELTM definuje třída `ELTMOperator`, která je potomkem obecné třídy `ToneMapping`. Na svém rozhraní nabízí metody pro nastavení svých konfigurovatelných parametrů a referenčního snímku pro danou sekvenci. Také implementuje metodu `ELTMOperator::map_frame` pro aplikaci mapování tónů na jeden vybraný snímek. Pomocí parametru této metody lze nastavit, zda mapování proběhne staticky či temporálně. Třída, která obsahuje z pohledu objektově orientovaného paradigmatu pouze statické metody, se nazývá `utils` a slouží pro pomocné maticové výpočty. Zdrojový kód je bohatě komentován, což umožňuje snadnou orientaci programátora při procházení zdrojových souborů aplikace.

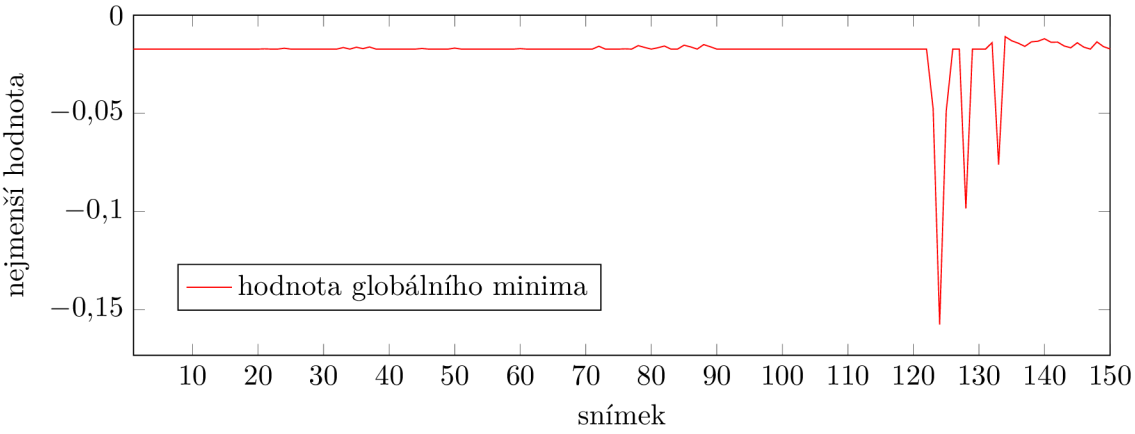
5.4.1 Temporální předzpracování snímků

Některé testovací sekvence mohou disponovat konstantním černým rámečkem okolo hran obrazu [32]. Aby rámeček neovlivňoval statistické hodnoty vstupního obrazu, je nutné detekovat jeho velikost a následně jej odstranit, což zajišťuje metoda `utils::remove_border`. Jestliže video opravdu rámeček obsahuje, je v úplném závěru zpracování daného snímku opět přidán zpět pomocí metody `utils::add_border`, aby byla zachována konzistence vůči vstupnímu obrazu.

[4] Testovací sada sekvencí [32]: https://hdr-2014.hdm-stuttgart.de/.

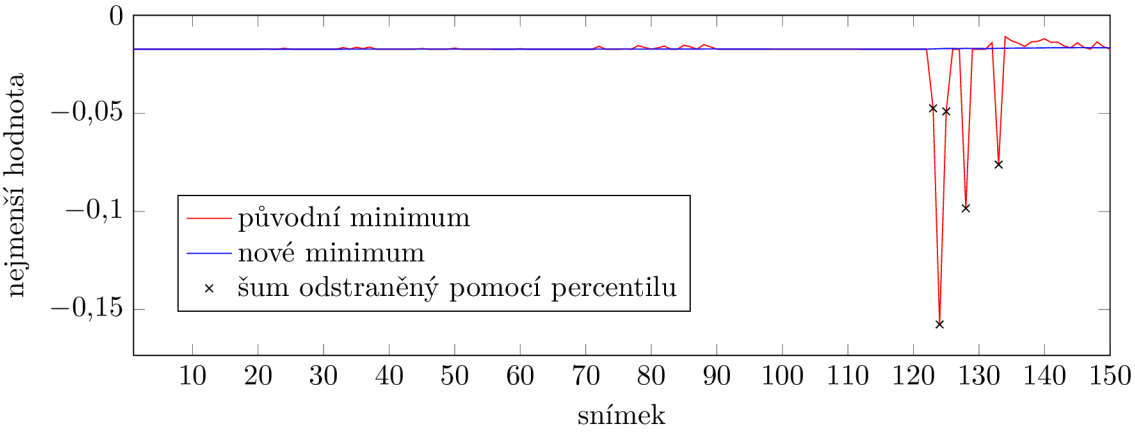
Vzhledem k faktu, že formát OpenEXR umožňuje pixelům nabývat i záporné hodnoty, je nutné snímky před aplikací operátoru ELTM předzpracovat. Určité testovací sekvence skutečně obsahují záporné hodnoty obrazových bodů, kdy se nejedná pouze o šum, ale také o relevantní obrazové informace. Je ale v podstatě nemožné přesně rozlišit, která data jsou relevantní a která představují nežádoucí šum. Naivním řešením by byl posun hodnot jednotlivých snímků do kladného rozsahu přičtením hodnoty globálního minima ke všem pixelům, nicméně tento přístup selhává v případě, kdy mezi snímky dojde ke skokové změně minima. Provedení posunu rozsahu by současně vedlo k nepřiměřenému posunu všech ostatních hodnot snímku. Ze stejného důvodu také není možné použít klasickou normalizaci jednotlivých snímků.

Pro řešení byla navržena metoda, kterou ilustruje následující příklad. Globální minimální hodnoty pixelů pro prvních 150 snímků z testovací sekvence *showgirl_02* [32] jsou znázorněny v následujícím grafu:



V grafu si lze povšimnout skokových změn globálních minimálních hodnot v temporální doméně, jmenovitě u snímků s pořadím 123, 124, 125, 128 a 133. Nelze říct, proč se hodnoty minim tak razantně mění, protože v samotné pořízené scéně k žádným změnám osvětlení nedochází. Zároveň ale není možné pro všechny snímky zcela zanedbat pixely se zápornou hodnotou, protože mohou obsahovat relevantní obrazová data.

Nejjednodušším řešením tohoto problému by jistě byl výpočet minimálních hodnot pomocí percentilu, kdy by došlo k vypuštění minimálních hodnot o velikosti menší vůči hodnotě stanovené pomocí percentilu. Tento přístup ovšem přináší dva problémy. Prvním z nich je výpočetní náročnost operace sloužící k výpočtu hodnoty pomocí percentilu, protože je nutné vždy seřadit všechny pixely vstupního snímku podle jejich velikosti. Druhým a podstatnějším problémem je ztráta obrazových dat s menší hodnotou, než je hranice stanovená percentilem. Oblasti snímků tvořené velmi malými hodnotami pixelů představují tmavá místa scény, na kterých se šum vizuálně projevuje nejvíce. Jestliže dochází k zanedbávání malých hodnot, roste také množství šumu v těchto místech. Za účelem řešení byla implementována metoda, která funguje na bázi temporálního prahu. Nežádoucí skokové změny minima jsou detekovány a odfiltrovány, a zároveň nedochází k téměř žádné ztrátě relevantních obrazových informací. Rozdíly mezi původními minimálními hodnotami a hodnotami minim po aplikaci temporálního předzpracování ilustruje následující graf:



Algoritmus předpokládá, že nejnižší možná hodnota konkrétního pixelu je rovna nule, která odpovídá absolutní černé barvě. Jestliže původní snímek obsahuje relevantní obrazová data se zápornými hodnotami, je nutné posunout celý rozsah snímku do rozsahu kladného. Principem je stanovení nové hodnoty globálního minima pro snímky se zápornými body, o jehož zápornou hodnotu je proveden posun celého vstupního rozsahu. Tento posun zajistí, že body předzpracovaného obrazu obsahují pouze kladné hodnoty. Nové minimum je vypočteno pomocí temporálního prahu, jehož hodnota představuje nové minimum. Jestliže v aktuálním vstupním snímku figuruje minimum s hodnotou menší než 0 a zároveň dojde k poklesu minima vůči hodnotě temporálního prahu o více než 2 %, proběhne výpočet nového minima pomocí percentilu o velikosti 0,005. Rozdíl je počítán jako procentuální rozdíl rozsahu daného původním minimem a průměrnou hodnotou obrazu, resp. percentilovým minimem a průměrnou hodnotou obrazu. Pokud je percentilové minimum stále menší než hodnota prahu, s velkou pravděpodobností se nejedná o změnu způsobenou šumem, ale minimum reálně kleslo, a je nutné snížit hodnotu přípustného prahu. V opačném případě algoritmus předpokládá, že propad minima byl způsobený nežádoucím šumem. Hladký pokles hodnoty prahu v čase je umožněn díky použití filtru s nekonečnou impulzní odezvou prvního řádu. Jestliže proběhl skutečný pokles minima, je nová hodnota prahu stanovena filtrem s konstantní impulzní odezvou o velikosti 0,10, jenž jako aktuální hodnotu uvažuje nové globální minimum vypočítané pomocí percentilu.

Výpočet hodnot na bázi percentilu zajišťuje metoda `utils::get_percentiles`, která nejprve z obrazových dat matice vytvoří matici s jedním řádkem, jehož prvky odpovídají všem bodům původního obrazu. Vzniká tedy jednorozměrná matice, která obsahuje všechna data a liší se pouze proporčně. Nyní lze na matici aplikovat metodu `cv::sort`, jež umožňuje vzestupně seřadit hodnoty jednotlivých řádků. Jelikož matice disponuje pouze jedním řádkem obsahujícím všechna obrazová data, po seřazení jeho hodnot lze získat hodnotu odpovídající požadovanému percentilu pomocí přístupu k prvku na určité pozici. Stanovení hodnoty přístupového indexu je dáno vynásobením velikosti percentilu a celkového počtu pixelů matice. Podmínkou pro správné fungování metody `cv::sort` je spojité uložení obrazových dat v paměti. K ověření této skutečnosti slouží metoda `cv::isContinuous`. Pakliže není matice v paměti uložena spojitě, je nutné data uložit do datového typu `std::vector` a následně na něj aplikovat řazení pomocí metody `std::sort` ze standardní knihovny jazyka C++. Toto řešení je ale přibližně třikrát až čtyřikrát pomalejší, než řazení pomocí metody knihovny OpenCV. Výpočet hodnoty dané percentilem probíhá v rámci celého algoritmu operátoru celkem čtyřikrát, pro snímky nesoucí záporná obrazová data a nekonečně velké hodnoty pixelů potenciálně až šestkrát. Použití metody `cv::sort` pro řazení prvků tedy

umožňuje razantní zrychlení celého algoritmu. Jelikož v dalších krocích výpočtu se hodnota na bázi percentilů počítá vždy dvakrát po sobě (pro minimum a maximum), implementovaná metoda `utils::get_percentiles` dokáže vypočítat dvě hodnoty najednou, a není tedy nutné řadit všechny body obrazu dvakrát.

Otázkou zůstává, zda je vhodné, aby hodnota minimálního prahu mohla v čase narůstat. Při teoretickém pohledu na tuto skutečnost je jasné, že hodnota prahu by měla v zásadě pouze klesat, nebo zůstávat konstantní, protože vyjadřuje globální temporální minimum rozsahu hodnot všech vstupních snímků. Jestliže vstupní snímky obsahují relevantní data se zápornými hodnotami, je možné, že v budoucnosti se snímky s podobnými hodnotami pixelů opět vyskytnou. Pokud by ale vstupní sekvence snímků obsahovala několik po sobě jdoucích snímků nesoucích velké množství šumu, který by se nepodařilo odfiltrovat pomocí percentilu, algoritmus už by se nikdy nedokázal zotavit. Vzhledem k faktu, že záporné hodnoty pixelů jsou v principu nežádoucí a vyskytují se pouze ve velmi omezené množině dostupných HDR sekvencí, algoritmus dovoluje, aby hodnota prahu mohla velmi pomalu konvergovat k nule v případě, že reálná globální minima snímků oproti prahu dlouhodobě stoupají. Použití filtru typu IIR prvního řádu s hodnotou parametru impulzní odezvy 0,01 umožňuje dosáhnout velmi pozvolného nárůstu hodnoty prahu vyjadřujícího spodní hranu využitelného vstupního rozsahu.

Dalšího zvýšení odolnosti algoritmu vůči neočekávaným datům lze docílit tak, že pro první snímek ze sekvence probíhá vždy výpočet minima pomocí percentilu. To platí samozřejmě pouze za předpokladu, že obraz obsahuje záporné hodnoty pixelů. Jestliže se rozsah od percentilového minima po průměrnou hodnotu obrazu liší vůči odpovídajícímu reálnému rozsahu o méně než 2 %, je počáteční práh nastaven na původní hodnotu vstupního minima. V opačném případě je pro výchozí hodnotu temporálního prahu použito minimum získané na základě percentilu. Stejný princip je také uplatněn pro předzpracování uživatelem vybraného referenčního snímku sekvence.

V zásadě tedy platí, že algoritmus předzpracování dokáže v případě výskytu záporných hodnot vyhodnotit, zda se jedná o relevantní data, nebo o nežádoucí šum. Šum je z obrazu následně odfiltrován a rozsah vstupních dat je posunut do kladných hodnot. Výpočet probíhá temporálně a vyhlazuje skokové změny globálních minimálních hodnot snímků, čímž je zvýšena celková odolnost operátoru vůči jasovým artefaktům.

5.4.2 Extrakce relativní luminance

Získání hodnot relativní luminance pomocí některé z metod implementovaných v rámci OpenCV není možné přímo, má-li výpočet hodnot odpovídat doporučení ITU-R BT. 709 (tedy rovnici 4.12). Jelikož bylo pro tento účel potřeba vytvořit vlastní implementaci, byl oproti původní variantě operátoru zvolen alternativní vzorec, který respektuje novější sadu doporučení ITU-R BT. 2100 (viz podkapitolu 3.2.1):

$$Y = 0.2627 \cdot R + 0.6780 \cdot G + 0.0593 \cdot B. \tag{5.1}$$

Aby nebylo nutné sekvenčně procházet celou vstupní obrazovou matici a aplikovat tuto operaci bod po bodu, byla implementována metoda pracující na procesoru paralelně. Knihovna OpenCV poskytuje obecnou třídu pro paralelní maticové výpočty s názvem `cv::ParallelLoopBody` a metodu `cv::parallel_for_`, které výrazně zvyšují míru abstrakce při návrhu paralelně pracujícího zdrojového kódu. Pro implementaci požadované operace nad celou maticí, kterou v tomto případě představuje výpočet relativní luminance, je nutné definovat pouze samotnou základní operaci a způsob, jak „rozdělit“ vstupní data

do jednotlivých vláken spouštěných paralelně. V rámci konkrétního vlákna probíhá klasický sekvenční výpočet po jednotlivých bodech pomocí definované operace. Při testování této metody vyšlo najevo, že pro snímky se šířkou 1920 a výškou 1080 pixelů představuje paralelní výpočet přibližně trojnásobné zrychlení oproti sekvenčnímu přístupu. Rychlost byla měřena s použitím procesoru podporujícího výpočet ve čtyřech vláknech.

Protože původní varianta operátoru neumožňuje pracovat se snímky obsahující pixely o nekonečné hodnotě a nedokáže se ani z této skutečnosti zotavit v čase, byla přidána funkcionalita řešící tento problém. Jestliže se ve vrstvě relativní luminance nekonečně velké hodnoty skutečně nacházejí, proběhne jejich vynucená saturace na maximální hodnotu nově stanoveného rozsahu. Horní hranice rozsahu je stanovena na hodnotu získanou pomocí percentilu o velikosti 99, přičemž do výpočtu nejsou zahrnuty body s nekonečně velkou hodnotou. Dojde tedy i k saturaci bodů tvořící jedno procento pixelů obrazu s nejvyššími hodnotami. Tato skutečnost nepředstavuje problém, protože saturované pixely představují absolutně bílé hodnoty a nedochází k produkci viditelného šumu.

5.4.3 Dekompozice obrazu a zpracování detailů

Před zahájením samotné dekompozice vrstvy relativní luminance je nutné provést její převod do logaritmické domény. Aby nedocházelo k problémům při logaritmizaci hodnot limitně se blížících nule, je nutné přičíst ke všem pixelům vrstvy konstantu o velikosti 10^{-6}.

Proces dekompozice implementuje metoda `ELTMOperator::separate_layers`. Vlastní extrakce základové a dvou detaily nesoucích vrstev probíhá pomocí filtrace řízeným filtrem, který podrobně rozebírá podkapitola 4.1.3. Knihovna OpenCV nabízí implementaci řízeného filtru v podobě metody `cv::ximgproc::guidedFilter`. Rozdělení na jednotlivé vrstvy obrazu je provedeno pomocí sekvence operací, jež popisuje série rovnic 4.14. Filtrace obrazu zde probíhá dvakrát.

Jako vstupní i řídící obraz pro první aplikaci filtru slouží vrstva relativní luminance převedená do logaritmické domény a poloměr konvolučního jádra filtru stanovuje uživatel. Velikost regualarizačního parametru filtru je nastavena pevně na konstantní hodnotu 0,1. Funkce clip figurující v původních rovnicích je implementována metodou `utils::clip_mat` s parametem `limit`, která pomocí metod knihovny OpenCV, `cv::min` a `cv:max`, saturuje hodnoty přesahující interval $[-limit, limit]$ jeho mezními hodnotami.

Vstupním i řídícím obrazem pro v pořadí druhou filtraci je dočasná základová vrstva tvořena vrstvou relativní luminance po odečtení jemných detailů. Odečtení detailů zprostředkovává metoda `cv::subtract`. Poloměr konvolučního jádra filtru pro extrakci hrubých detailů je vypočten jako jedna desetina kratšího rozměru vstupního obrazu. Výstupem filtru je rovina hrubých detailů a jejím odečtením od dočasné základové vrstvy vzniká konečná logaritmická základová vrstva nesoucí osvětlení scény, na kterou lze v pozdější fázi aplikovat kompresi tonality.

Pro posílení detailů snímku je vypočtena proměnlivá mapa zesílení pomocí metody `cv::multiply`, která vynásobí všechny obrazové body základové vrstvy hodnotou $-0,4$. Následuje saturace hodnot získané matice, které mají velikost menší než 1, pomocí prahování metodou `cv::max`. V posledním kroku zpracování detailů jsou obě vrstvy detailů vynásobeny získanou mapou zesílení a následně sečteny. Vzniká tak jedna matice obsahující logaritmická data, která reprezentuje hrubé i jemné detaily obrazu.

5.4.4 Komprese rozsahu základové vrstvy

Statistické hodnoty α, β z rovnic 4.16 jsou získány ze základové vrstvy osvětlení na základě percentilu. Při volbě velikosti percentilů byly informace čerpány z článků věnujících se operátoru ELTM a jeho temporální adaptaci [27, 28]. Zvoleny byly vždy hodnoty bližší příslušné hranici rozsahu, konkrétně 0,01 pro minimum a 99,99 pro maximum, aby nedocházelo ke zbytečné ztrátě obrazových informací.

Jak již bylo zmíněno v teoretické části textu věnující se operátoru ELTM, komprese tónů je provedena pouze pro základovou vrstvu. Nejprve je nutný posun a škálování jejího rozsahu do předem stanoveného intervalu v logaritmické doméně. Jako první je proveden posun pomocí metody `cv::add`, která přičte ke všem prvkům statistickou hodnotu β. Škálování zajišťuje metoda `cv::multiply`, která násobí všechny prvky získaným koeficientem α. Za účelem faktické komprese tónů je nutné základovou vrstvu s posunutým rozsahem převést do lineární domény, k čemuž slouží metoda `utils::inverse_log_2` pracující na procesoru paralelně. Dále proběhne výpočet minimální a maximální hodnoty linearizované vrstvy, jež jsou opět stanoveny pomocí percentilu. Minimální hodnota je dána percentilem o hodnotě 0,1 a maximální percentilem o velikosti 99,9. Pro výpočet odpovídajících obrazových hodnot je opět použita již představená metoda `utils::get_percentiles`. Pro samotnou kompresi tonality slouží metoda `ELTMOperator::tone_compression`, která na základovou vrstvu postupně aplikuje maticové operace na základě rovnice 4.22.

5.4.5 Obnovení barev

Před samotným obnovením barev z původního HDR obrazu je nutné sloučit detaily nesoucí a komprimovanou základovou vrstvu osvětlení pomocí vzájemného vynásobení jejich prvků. Základová vrstva po kompresi tónů obsahuje lineární hodnoty, zatímco kombinovaná vrstva detailů je v této fázi stále uložena v logaritmické doméně a je nutné ji před sloučením linearizovat pomocí metody `utils::inverse_log_2`. Obnovení původních barev implementuje metoda `ELTMOperator::color_restoration`, jež nejprve rozdělí vstupní snímek s vysokým dynamickým rozsahem na jednotlivé barevné kanály. Pro každý barevný kanál následně proběhne obnovení barev podle vzorce 4.24, které umožňuje zvolit úroveň barevné saturace výsledného obrazu. V závěru jsou sloučeny vypočítané kanály a výsledná matice je převedena z typu `CV_32FC3` na typ `CV_8UC3`, jenž odpovídá klasickému formátu obrazových dat s nízkým, resp. běžným dynamickým rozsahem. Jestliže původní vstupní HDR snímek obsahoval rámeček okolo relevantních obrazových dat, je o něj výstupní snímek opět doplněn metodou `utils::add_border`.

5.4.6 Temporální zpracování snímků a zápis videa

Třída `ELTMOperator` uchovává proměnné statistické parametry pro temporální zpracování jako své třídní atributy a disponuje metodou simulující temporální výpočetní blok (viz podkapitolu 4.3.2). Tato metoda, nazvaná `ELTMOperator::temporal_iir_compute`, provádí pro každý statistický parametr výpočet jeho nové hodnoty na základě statistiky získané z aktuálního snímku a hodnoty aplikované pro předchozí snímek ze sekvence podle rovnice 4.25. Koeficient adaptace, vyjádřený ve zmíněné rovnici proměnnou k, odpovídá impulzní odezvě filtru typu IIR prvního řádu a udává míru adaptace obecného statistického parametru operátoru v čase pro jednotlivé snímky. Vysoké hodnoty koeficientu k znamenají vyšší míru adaptace příslušného parametru, ale potenciálně mohou způsobovat vznik temporálních artefaktů. Článek popisující temporální fungování operátoru ELTM [28] neuvádí,

jaké hodnoty koeficientů odezvy je vhodné pro jednotlivé parametry zvolit. Tato diplomová práce navrhuje, aby koeficienty pro jednotlivé parametry byly konfigurovatelné uživatelem, protože pro různé vstupní sekvence snímků lze dosáhnout odlišné míry temporální adaptace. Pakliže by koeficienty byly nastaveny na stejné hodnoty pro všechny vstupní sekvence, nebylo by pro některá videa možné dosáhnout maximální možné míry adaptace.

Aby nedošlo k nepřiměřenému nárůstu komplexity rozhraní operátoru, může uživatel konfigurovat míru adaptace velmi jednoduše prostřednictvím jednoho souhrnného parametru. Výsledné hodnoty koeficientů filtrů pro jednotlivé parametry jsou vypočteny podle následujícího vzorce 5.2:

$$k = v_a \cdot k_{max}, \tag{5.2}$$

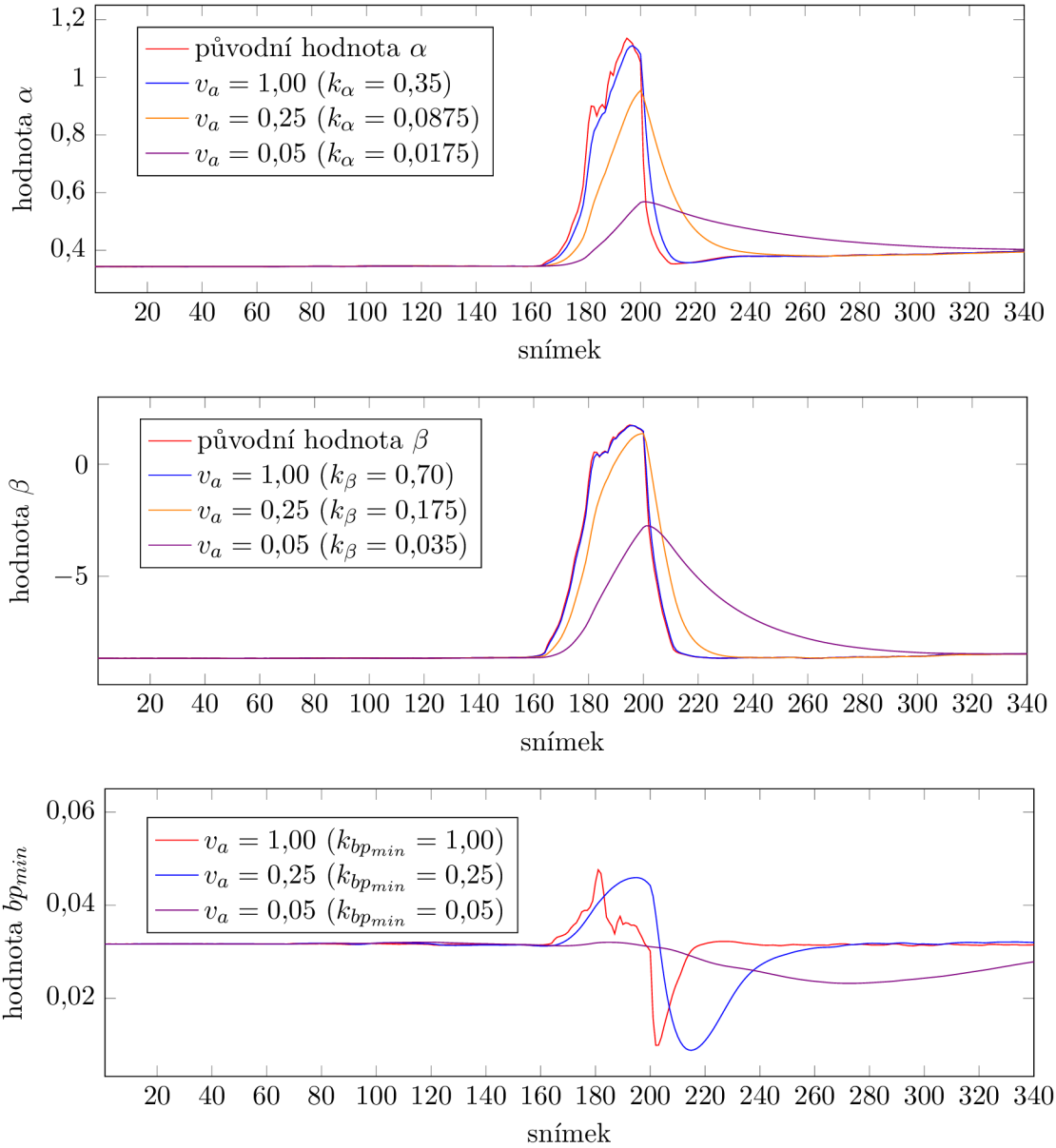
kde proměnná v_a vyjadřuje uživatelem konfigurovatelný parametr s rozsahem hodnot $[0, 1]$. Nízké hodnoty tohoto parametru znamenají, že operátor mapování tónů bude na změny ve scéně reagovat pomalu, čímž je zvýšena koherence výsledného videa. Opačný případ nastává při zvolení vysoké hodnoty, kdy probíhá rychlejší adaptace za cenu ztráty koherence videa. Každé vstupní video je v tomto ohledu unikátní a nelze před zahájením mapování tónu určit, zda dojde ve výsledné sekvenci k výskytu artefaktů. Z tohoto důvodu je výchozí hodnota parametru pro rychlost adaptace nastavena spíše konzervativně na velikost 0,2. Hodnota k_{max} z předešlé rovnice 5.2 představuje maximální povolenou hodnotu koeficientu k, jenž vyjadřuje reálně použitou velikost odezvy filtru pro konkrétní parametr. Maximální hodnoty koeficientů konkrétních parametrů byly stanoveny na základě matematického pohledu na fungování operátoru a rozsáhlého empirického testování. Stanovené hodnoty k_{max} jsou uvedeny v následující tabulce 5.1:

parametr	$\boldsymbol{k_{max}}$
α	0,35
β	0,70
BP_{min}	1,0
BP_{cmax}	0,35

Tabulka 5.1: Maximální hodnoty koeficientů filtrů pro temporální adaptaci statistik.

Vzhledem k tomu, že kontrolu všech uvedených koeficientů provádí uživatel přes jeden parametr, bylo důležité stanovit jejich maximální možnou velikost za základě jejich adekvátních poměrů. Statistická hodnota α slouží ke škálování celého vstupního rozsahu a je nutné, aby škálování probíhalo pomaleji než posun rozsahu, který přímo ovlivňuje proměnná β. Při příliš rychlé adaptaci hodnoty α může docházet ke ztrátě globální koherence kontrastu. Nejlepších výsledků se pro rozsáhlou testovací sadu videí podařilo dosáhnout při použití poloviční velikosti pro koeficient statistiky α oproti velikosti odezvy filtru pro posun rozsahu daný hodnotou β. Hodnota adaptačního koeficientu pro statistiku BP_{min} byla ponechána na maximální možné velikosti, tedy 1, a to ze dvou důvodů. Prvním důvodem je, že hodnota samotného parametru se počítá na bázi percentilu, který zabraňuje neočekávaným skokovým změnám parametru v čase. Druhým důvodem je fakt, že hodnoty α a β figurují v algoritmu před výpočtem hodnoty BP_{min}, a ovlivňují tak její velikost. To v principu znamená, že pro parametr BP_{min} probíhá nepřímá temporální adaptace ještě před samotnou aplikací filtru na jeho vlastní hodnotu. Jelikož statistika BP_{cmax} slouží podobně jako proměnná α ke škálování rozsahu hodnot, je pro jeho adaptační koeficient zvolena ekvivalentní maximální hodnota, tedy 0,35.

Následují grafy, které vyjadřují vliv parametru rychlosti adaptace na jednotlivé statistické hodnoty. Na vlastnostech výsledného videa se typicky nejvíce podílejí statistické proměnné α, β a BP_{min}. Pro každou z těchto proměnných je demonstrována míra „vyhlazení“ jejich hodnot v čase pomocí filtru s nekonečnou impulzní odezvou prvního řádu. Jako referenční video slouží série *showgirl_02* [32]. Video zobrazuje dívku sedící u stolu na pódiu divadelního sálu. Scéna je přibližně 160 prvních snímků osvětlena pomocí několika zdrojů světla. Poté, co dívka vstane od stolu a postaví se čelem ke kameře, všechny zdroje světla ve scéně zhasínají. Přibližně 40 dalších snímků v pořadí snímků zobrazuje velmi temnou scénu. Poté je na pódium vrženo velmi silné světlo pomocí reflektoru a dochází k velmi rapidní změně osvětlení celé scény. Reflektor zůstává nadále aktivní po celý zbytek sekvence.



Pro demonstraci rychlosti adaptace jednotlivých statistik byly pro parametr rychlosti adaptace v_a zvoleny hodnoty 1, 0,25 a 0,05. Jak už bylo zmíněno, vyšší hodnota umožňuje rychlejší adaptaci statistických hodnot. Výpočty velikosti koeficientů odpovídají rovnici 5.2.

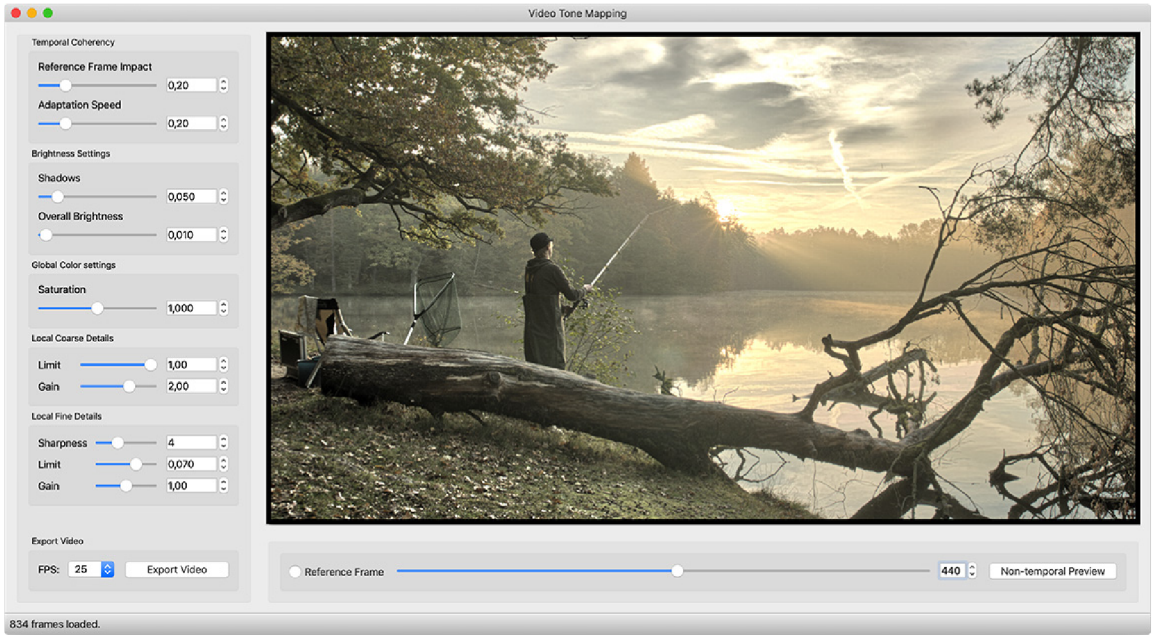
Z grafů je patrné, že pro tuto konkrétní sekvenci lze dosáhnout koherentních výstupů, jež neobsahující skokové změny statistických hodnot mezi snímky i při poměrně vysokých hodnotách parametru pro konfiguraci adaptační rychlosti.

V některých případech je ale žádoucí reagovat na změny osvětlení pomaleji. Požadavek na pomalejší adaptaci může přijít v případě použití snímků ze stacionárních dohledových kamer, nebo například sekvencí, u kterých uživatel předem ví, že ke změnám osvětlení bude doházet pouze pozvolna.

Míru koherence videa vůči zvolenému referenčnímu snímku lze dodatečně konfigurovat parametrem, který taktéž může nabývat hodnot 0 až 1. Nejvyšší hodnota znamená maximální koherenci videa vůči vybranému snímku. Oproti původnímu návrhu operátoru ELTM má tedy tento parametr opačný vliv. Jeho hodnota ovlivňuje pouze proměnnou β a na ostatní statistiky nemá vliv. Výpočet temporálně adaptované statistické hodnoty β_A probíhá stejně jako v případě ostatních statistik pomocí filtru typu IIR prvního řádu. Až před její samotnou aplikací dojde k přepočtu vůči referenčnímu snímku na hodnotu β_{Aref} (viz rovnice 4.25, 4.26). Operátor pro další snímek v pořadí neuvažuje hodnotu β_{Aref} jako statistickou, k tomuto účelu slouží původní adaptovaný parametr β_A před provedením jeho přepočítání vůči referenčnímu obrazu.

5.5 Grafické uživatelské rozhraní

Podoba grafického uživatelského rozhraní aplikace vychází z běžných editorů pro úpravu videa a obrázků. Dominantním prvkem je oblast pro náhled zvoleného snímku videa, pod kterým se nachází ovládací prvky pro volbu snímku sloužícího k náhledu. Je možné zvolit určitý snímek pomocí posuvného prvku, nebo zadat jeho pořadové číslo do textového pole explicitně. Nalevo od popsaných částí se nachází hlavní kontrolní panel pro ovládání operátoru mapování tónů. Jednotlivé ovládací prvky jsou rozděleny sémanticky podle svého vlivu na výsledný obraz do dílčích celků. Implementované grafické rozhraní demonstruje obrázek 5.1.



Obrázek 5.1: Grafické rozhraní aplikace. Snímek pochází z videa *fishing_longshot* [32].

Nejvýše se v ovládacím panelu nachází elementy pro kontrolu temporálních parametrů. Uživatel zde může nastavit, jak velký vliv bude referenční snímek mít na ostatní snímky videa a rychlost temporální adaptace operátoru. Následující blok prvků obsahuje položky pro úpravu parametrů tónové komprese. První element odpovídá parametru BP_{cmin} pro nastavení spodní hranice výstupního dynamického rozsahu videa a druhý parametru p, který ovlivňuje sílu komprese tónů (viz sekci 4.5.6). Přestože poslední dva zmíněné prvky slouží pro konfiguraci komprese tonality, v rámci grafického rozhraní jsou označeny jako parametry pro nastavení jasu, protože úprava jejich hodnot se na výsledném videu nejvíce projeví právě změnou jasu obrazu. Další blok v pořadí obsahuje pouze jeden element, skrze který může uživatel regulovat barevnou saturaci obrazu.

Následují dvě skupiny prvků pro kontrolu hrubých a jemných detailů obrazu. V obou případech lze nastavit práh pro jejich maximální hodnotu (viz sérii rovnic 4.14) a míru zesílení jejich kontrastu popsanou v podkapitole 4.5.5. Uživatel také může zvolit velikost konvolučního jádra řízeného filtru, jenž slouží pro extrakci vrstvy jemných detailů, pomocí separátního elementu.

Poslední blok hlavního kontrolního panelu obsahuje prvek pro volbu snímkové frekvence výsledného videa. Výchozí hodnota je nastavena na velikost 25. Uživatel může pomocí kombinovaného seznamu volit mezi frekvencemi 25, 30, 60 a 120 snímků za sekundu. Po zahájení mapování tónů videa se uživateli zobrazí modální okno ukazující aktuální fázi probíhajícího výpočtu. Pomocí tlačítka, které se nachází v tomto okně, může uživatel přerušit probíhající mapování tonality. Jelikož ukládání snímků do videa probíhá sekvenčně, je možné vytvořit video obsahující snímky, které již byly před přerušením zpracovány, a nedochází tedy ke ztrátě výstupních dat.

Pro zvýšení pohodlí uživatele byla implementována možnost přiblížit a oddálit snímek zobrazený v náhledu. Lze také přizpůsobit velikost zobrazeného snímku ploše dostupné pro náhled, což může být užitečné například po změně velikosti okna aplikace. Je nutné poznamenat, že pro snímek sloužící k náhledu neprobíhá temporální adaptace operátoru v čase a je ovlivněn pouze referenčním snímkem. Ve velké většině situací ale obraz disponuje takřka shodnými vlastnostmi, jako daný snímek ve výstupním videu po aplikaci temporálního mapování tónů.

5.5.1 Implementace grafického uživatelského rozhraní

Jak již bylo zmíněno v úvodu této kapitoly, pro vývoj grafického uživatelského rozhraní byl vybrán framework Qt. Hlavní okno aplikace je realizováno za použití instance třídy `MainWindow`, která je potomkem obecné třídy Qt s názvem `QMainWindow` a obsahuje veškeré prvky primárního okna. Dialogová okna pro načítání a export souborů zprostředkovává třída `QFileDialog` a modální okno zobrazující aktuální stav probíhajícího temporálního mapování je realizováno pomocí třídy `QProgressDialog`.

Po spuštění aplikace dojde k okamžité inicializaci hlavního okna a jeho prvků pomocí metody `MainWindow::init_gui`, která vytvoří a následně patřičně propojí spolu související elementy. Po zdařilé inicializaci aplikace očekává, že uživatel zvolí ve hlavní nabídce možnost pro načtení vstupních snímků a vybere požadovanou sekvenci. Načítání snímků a inicializaci náhledu zajišťuje metoda `MainWindow::load_hdr_frames`, jež po načtení snímků vytvoří instanci operátoru mapování tónů (třídy `ELTMOperator`) pro účel náhledu a zpřístupní uživateli ovládací prvky rozhraní. V další fázi může uživatel měnit hodnoty parametrů operátoru a zkoumat jejich vliv na výsledný obraz stiskem tlačítka pro zobrazení náhledu. Vykreslení aktuálního náhledu má na starost metoda `MainWindow::show_preview`, která

spustí statické mapování tónů pro vybraný snímek a převede výstupní obrazovou matici operátoru z datového typu `cv::Mat` na typ `QImage`. Framework Qt nedokáže pracovat s datovými typy definovanými knihovnou OpenCV, a proto je nutné provést konverzi.

Uživatel má kromě exportu videa také možnost uložit samostatný statický snímek. Uložen je vždy ten snímek, který je aktuálně zobrazen v náhledu, a uživatel může vybírat mezi formáty JPEG (*.jpg), PNG (*.png), BMP (*.bmp) a TIFF (*.tiff). Za vytváření výstupních statických obrázků odpovídá framework Qt prostřednictvím instance třídy `QImageWriter`. Export temporálně mapovaného videa proběhne po stisknutí příslušného tlačítka a realizuje jej metoda `MainWindow::tonemap_video`, která po získání cesty pro výstupní soubor vytvoří novou instanci třídy `ELTMOperator` a iterativně volá jeho metodu `ELTMOperator::map_frame`. Zápis jednotlivých snímků do výstupního videa je řešen pomocí knihovny OpenCV s použitím instance třídy `cv::VideoWriter`, jež je atributem obecné třídy pro mapování tónů zvané `ToneMapping`. Přidávání snímků do konečného videa probíhá v reálném čase. Po dokončení zpracování konkrétního snímku proběhne jeho zápis do video souboru prostřednictvím metody `ToneMapping::append_frame_to_video`, která jej připojí na konec videa.

Kapitola 6

Ukázka a zhodnocení výsledků

Tato kapitola nejprve demonstruje vliv jednotlivých uživatelských parametrů na výsledné sekvence. Porovnání obrazu získaného pomocí různých hodnot proměnných pro statické snímky je následováno ukázkou vlivu temporálních parametrů na výsledné video. Poté jsou produkované výstupy aplikace vizuálně srovnány s výsledky jiných operátorů mapování tónů. Na závěr této kapitoly je stručně uvedena rychlost implementovaného algoritmu.

Tabulka 6.1 uvádí výchozí hodnoty jednotlivých uživatelských parametrů a přípustný rozsah, ze kterého může uživatel jejich hodnoty vybírat. Výchozí hodnoty byly zvoleny tak, aby operátor v základním nastavení poskytoval nejlepší možné výsledky pro co největší množství vstupních dat.

název parametru	**proměnná (z rovnice)**	**povolený rozsah**
vliv referenčního snímku	Ψ (4.26)	[0; 1] (0,2)
rychlost temporální adaptace	v_a (5.2)	[0; 1] (0,2)
stíny	BP_{cmin} (4.22)	[0; 0,4] (0,08)
celkový jas	p (4.22)	[0,001; 0,5] (0,03)
barevná saturace	s (4.24)	[0; 2] (1)
limit hrubých detailů	λ_C (4.14)	[0; 1] (1)
zesílení hrubých detailů	η_C (4.19)	[0; 3] (1,5)
rozlišení jemných detailů	r_s (4.14)	[0; 10] (3)
limit jemných detailů	λ_F (4.14)	[0; 0,1] (0,02)
zesílení jemných detailů	η_F (4.19)	[0; 2] (1)

Tabulka 6.1: Zasazení uživatelských parametrů do teoretického kontextu. Jejich výchozí hodnoty jsou uvedeny v kulatých závorkách ve sloupci, který definuje možné rozsahy hodnot pro jednotlivé položky.

6.1 Dopad statických parametrů na výsledný obraz

Následující obrázky 6.1 slouží pro vizualizaci vlivu nejzásadnějších statických parametrů na konečné snímky. Kromě modifikace vždy uvedeného parametru byly pro všechny snímky použity výchozí hodnoty ostatních proměnných. Pouze v případě jemných detailů η_F byl zvýšen práh jejich limitu λ_F, aby byl dopad parametru zřetelnější.

$BP_{cmin} = 0$ $BP_{cmin} = 0{,}2$ $BP_{cmin} = 0{,}4$

$p = 0{,}001$ $p = 0{,}05$ $p = 0{,}5$

$s = 0$ $s = 1$ $s = 2$

$\eta_C = 0$ $\eta_C = 1{,}5$ $\eta_C = 3$

$\eta_F = 0; \lambda_F = 1$ $\eta_F = 1; \lambda_F = 1$ $\eta_F = 2; \lambda_F = 1$

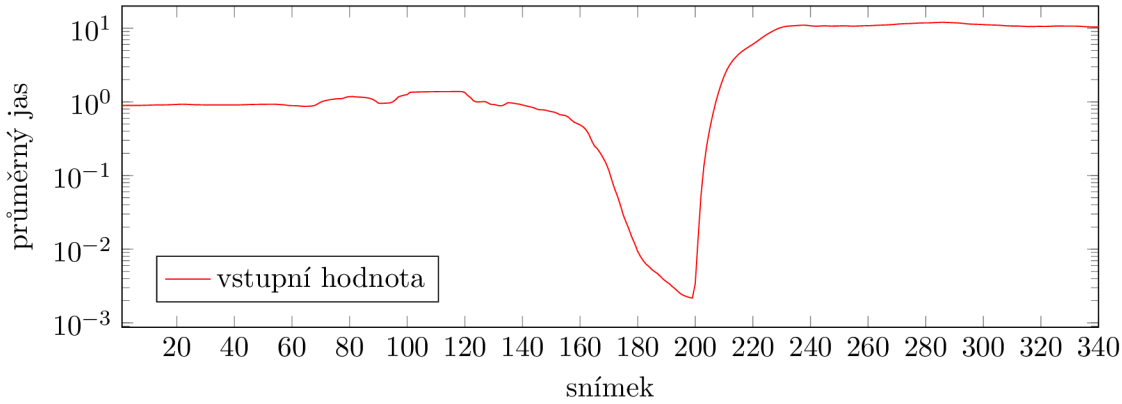
Obrázek 6.1: Dopad vybraných parametrů na výsledný obraz[1].

Kombinací různých hodnot uživatelských parametrů lze dosáhnout velmi rozmanitých výsledků. Je-li cílem vyvolat umělecký dojem, lze zesílit kontrast detailů a zvýšit barevnou saturaci. Pokud je motivací co nejvěrněji reprezentovat scénu, je vhodné nastavit parametry konzervativněji, na základě vlastností vstupního obrazu. Velmi jednoduše lze získat také černobílé snímky, jejich vytvoření je možné pomocí úplného potlačení barevné saturace.

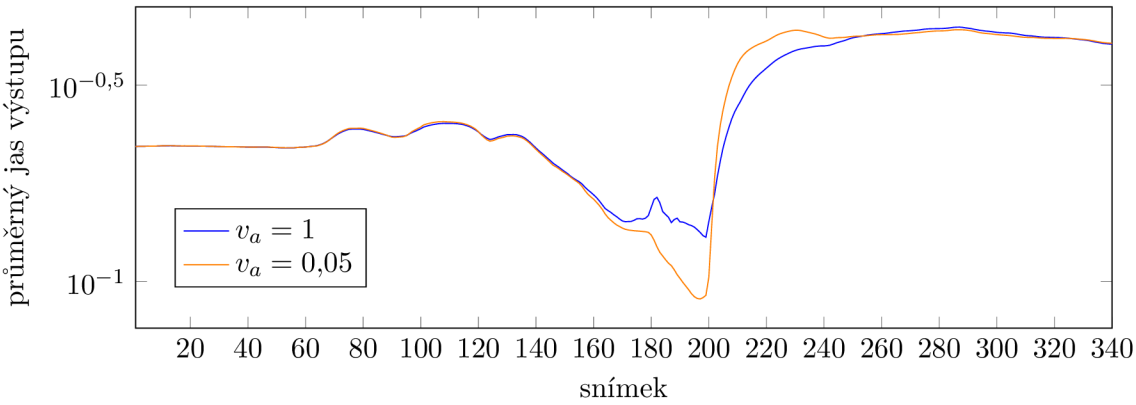
[1]Tištěná verze tohoto textu může obrázky oproti jejich zobrazení na displeji zkreslovat.

6.2 Ukázka vlivu rychlosti temporální adaptace

Výstupy algoritmu v podobě statických obrázků ukazují, že pro libovolný snímek lze dosáhnout výborného výsledku. Temporální výpočet musí zajistit, aby po sobě jdoucí snímky zůstaly vzájemně koherentní a aby video neobsahovalo obrazové artefakty. Jelikož mapování tónů v časové doméně probíhá sekvenčně snímek po snímku, kdy algoritmus nepočítá přímo s obrazovými daty okolních snímků, ale pouze s globálními statistickými hodnotami, je vyloučen vznik pohybových artefaktů. Jasové artefakty mohou být efektivně potlačeny díky zvolení vhodné hodnoty parametru pro rychlost temporální adaptace. Nejlépe lze vyjádřit výskyt problikávání pomocí grafů reprezentujících průměrné hodnoty jasu snímků před a po aplikaci temporálního mapování tónů. Nejprve je uveden graf průměrných hodnot jasu vstupní HDR sekvence *showgirl_02* [32] v logaritmickém měřítku:



Jestliže nedochází ve stupních datech ke skokovým změnám jasu, nesmí k nim docházet ani ve výstupním videu. Pro demonstraci dosažitelných výsledků byly vybrány jako hodnoty parametru adaptační rychlosti konstanty 1 a 0,05. Následuje graf ukazující výsledné hodnoty průměrného jasu pro obě dvě varianty:



V případě nastavení rychlosti adaptace na hodnotu 1 lze pozorovat neočekávaný vzestup průměrné hodnoty jasu okolo 180. snímku. Tato skutečnost je způsobena tím, že algoritmus cílí na co nejefektivnější využití celého výstupního dynamického rozsahu. V této fázi sekvence dochází k vzestupu statistické hodnoty bp_{min}, který lze pozorovat ve třetím grafu podkapitoly 5.4.6. Není jasné, proč tato hodnota stoupá, když dochází k poklesu celkového osvětlení scény. Podobné nepředvídatelné situace lze řešit snížením rychlosti adaptace

algoritmu. Jestliže adaptace probíhá pomaleji, je zaručena vyšší robustnost vůči jasovým artefaktům. Problémem pomalé adaptace je ale delší reakční doba po dramatické změně osvětlení scény. Algoritmu totiž trvá delší dobu, než se hodnoty jeho statistických parametrů ustálí vůči odpovídajícím hodnotám v původní sekvenci. Může se pak stát, že celkový jas koreluje s původní scénou, ale může docházet k vypouštění hodnot z okrajů dynamického rozsahu původního snímku. Tento proces lze ilustrovat na příkladu. Po rozsvícení reflektoru ve zmíněné scéně (ve snímku č. 200) sice průměrná hodnota jasu dobře koreluje se vstupními daty, nicméně dochází ke ztrátě detailů ve světlých a tmavých částech snímků, protože algoritmus reaguje na změnu osvětlení delší dobu. Během této doby nedochází k optimálnímu využití výstupního dynamického rozsahu. Následují obrázky popisující tuto skutečnost.

$v_a = 1$

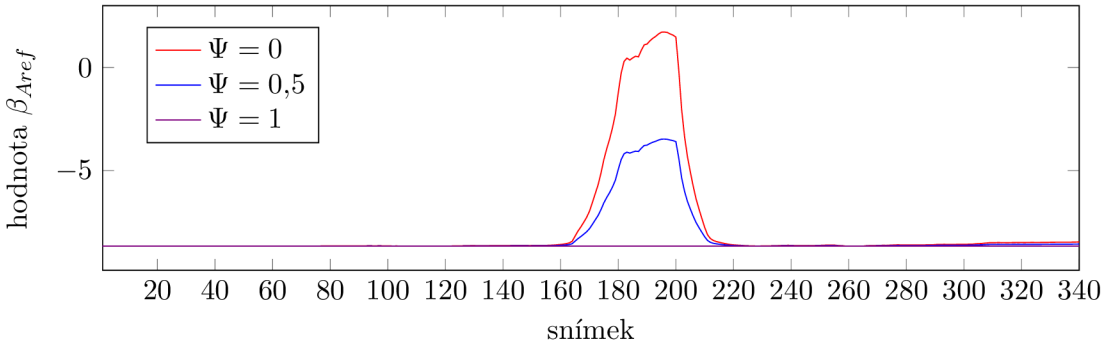
$v_a = 0{,}05$

Obrázek 6.2: Ukázka vlivu adaptační rychlosti na výsledné video. Horní i spodní sekvence čtyř obrázků zobrazují snímky č. 199, 201, 205 a 210 z videa *showgirl_02* [32]. Ve druhé situaci dochází k pomalejší adaptaci dynamického rozsahu a po rozsvícení reflektoru může na krátkou chvíli výsledné video působit dojmem, že bylo pořízeno s příliš vysokou expozicí.

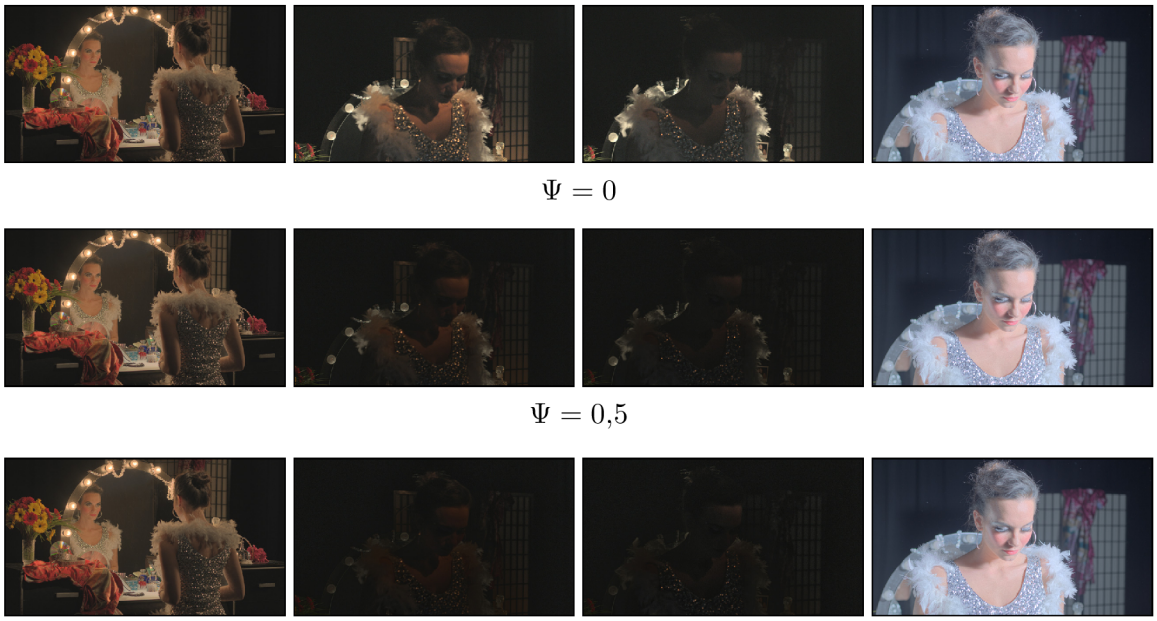
Série obrázků 6.2 ukazují, že při zvolení vysoké hodnoty adaptační rychlosti využívají jednotlivé snímky lépe potenciál dynamického rozsahu obrazu. Menší hodnoty parametru zajišťují větší potlačení jasových artefaktů, ale snímky mohou ztrácet detaily ve světlých a tmavých oblastech. Je na uživateli, aby zvolil vhodnou rychlost adaptace pro konkrétní vstupní video.

6.3 Vliv referenčního snímku na výsledné video

Jak velký vliv bude mít referenční snímek na zbytek videa, může uživatel ovlivnit pomocí dedikovaného parametru. Jedinou statistickou hodnotou, kterou tento parametr ovlivňuje, je proměnná β odpovídající za posun rozsahu základové vrstvy osvětlení. Jestliže je kontrolní parametr vlivu nastaven na nulovou hodnotu, použije se posun rozsahu vypočítaný pouze na základě konkrétního snímku a zvolené rychlosti temporální adaptace. Vyšší hodnoty znamenají, že velikost posunu bude stanovena pomocí váženého průměru hodnoty vypočítané pro daný snímek a hodnoty posunu získané z referenčního snímku (viz rovnici 4.26). Koeficient váhy přímo odpovídá uživatelskému parametru určujícího míru vlivu referenčního obrazu, přičemž větší hodnoty znamenají vyšší váhu hodnoty posunu β_{ref} získané z referenčního snímku. Tento princip ilustruje následující graf:



Na grafu jsou vyobrazeny hodnoty β_{Aref} pro celou sekvenci *showgirl_02* [32]. Pro vyšší hodnoty uživatelského parametru vlivu referenčního snímku (tedy Ψ z rov. 4.26) platí, že ve výsledném videu jsou lépe zachovány původní jasové vlastnosti reálné scény, protože pro jednotlivé snímky neprobíhá posun dynamického rozsahu pouze na základě jejich vlastního rozsahu, ale je zohledněn i obraz referenční. S rostoucí velikostí parametru se změny osvětlení scény projevují ve výsledném videu razantněji. Parametr tedy umožňuje zvolit poměr mezi utilizací dynamického rozsahu a zachováním globálních jasových vlastností původní scény. Přínosem je tedy možnost modifikace temporálních nejen jasových, ale i kontrastních vlastností videa podle požadavků uživatele. Vliv parametru je opět demonstrován na stejné sekvenci jako v případě předchozího grafu:



Obrázek 6.3: Tři sekvence ukazující vliv referenčního snímku na video.

Sekvence obrázků 6.3 začínají vždy snímkem č. 1, který byl zvolen jako referenční. Další tři obrázky pokaždé ukazují snímky č. 185, 198 a 225. Lze si všimnout, že u snímků 185 a 198 byl při nastavení vyššího vlivu referenčního snímku potlačen globální jas a kontrast. Snímek 225 disponuje podobnou hodnotou globálního maxima jako referenční obraz, a proto došlo ve všech sekvencích k výrazně menšímu posunu rozsahu oproti snímkům 185 a 198.

6.4 Ukázka výsledků a srovnání s nejvyspělejšími operátory

Tato podkapitola ilustruje dosažené výsledky a obsahuje velké množství obrázků, které byly získány z výsledných videí. Nejprve jsou na vybraných sekvencích vizuálně demonstrovány vlastnosti výstupního obrazu a následuje srovnání výsledných snímků s nejvyspělejšími existujícími operátory pro temporální mapování tónů.

desert [8, 9]

showgirl_02 [32]

fishing_longshot [32]

smith_hammering [32]

hallway [26]

poker_fullshot [32]

Obrázek 6.4: Aplikace dosahuje skvělých výsledků pro širokou škálu způsobů osvětlení scény.

Porovnávání kvality výsledků jednotlivých operátorů pro mapování tónů je velmi komplexní záležitost. Existuje mnoho různých přístupů pro srovnávání, které přesahují do oborů jako jsou neurologie, fyziologie vidění, psychologie a mnoha dalších [13, 14, 15]. Vizuální dojem z obrazu je navíc značně subjektivní. Při srovnání s existujícími operátory je ale zřejmé, že vlastnosti výsledného obrazu jsou skutečně výborné a implementaci lze poměřovat s nejvyspělejšími algoritmy. Referenční výstupy ostatních operátorů byly získány z doprovodných materiálů článku popisujícího temporální operátor Noise-aware TMO [25].

implementovaná aplikace

Mal-adaptation TMO [41]

Display Adaptive TMO [54]

Virtual Exposures TMO [6]

Temporal Coherence TMO [7]

Zonal Temporal Coherence TMO [9]

Motion Path Filtering TMO [4]

Noise-aware TMO [25]

Obrázek 6.5: Vizuální srovnání různých operátorů pro sekvenci *car_fullshot* [32]. Na výstupním obrazu implementované aplikace lze pozorovat velmi dobrý kontrast na globální i lokální úrovni. Zároveň dochází oproti ostatním operátorům k menšímu potlačování lokálních detailů, čehož si lze všimnout na dlaždicích ve scéně. Je nutné poznamenat, že rozdíly nemusí být v tištěné verzi textu oproti zobrazení na displeji tak dobře patrné.

implementovaná aplikace

Mal-adaptation TMO [41]

Display Adaptive TMO [54]

Virtual Exposures TMO [6]

Temporal Coherence TMO [7]

Zonal Temporal Coherence TMO [9]

Motion Path Filtering TMO [4]

Noise-aware TMO [25]

Obrázek 6.6: Výsledky mapování videosekvence *smith_hammering* [32] ukazují, jak dokáží jednotlivé operátory redukovat šum ve tmavých částech scény. Výsledky implementované aplikace obstávají velmi dobře a mohou konkurovat i algoritmům, které šum explicitně zohledňují [6, 25]. Zajímavé je také sledovat, jak jednotlivé operátory interpretují oblast plamenů ohně.

implementovaná aplikace

Mal-adaptation TMO [41]

Display Adaptive TMO [54]

Virtual Exposures TMO [6]

Temporal Coherence TMO [7]

Zonal Temporal Coherence TMO [9]

Motion Path Filtering TMO [4]

Noise-aware TMO [25]

Obrázek 6.7: Na výsledcích pro sekvenci *showgirl_02* [32] lze pozorovat u obrazu získaného pomocí implementované aplikace kromě skvělého globálního kontrastu také velmi dobré zachování jemných detailů, zejména v oblasti zrcadla, kde jsou viditelné i drobné nečistoty na skle. Zároveň nedochází k nadměrnému projevu šumu a jsou zachovány i detaily závěsu, který je ve scéně nejdále od kamery.

6.5 Rychlost implementované aplikace

Tato práce neklade zvláštní důraz na optimalizaci rychlosti vyvinutého programu a spíše se věnuje problematice jeho samotné funkcionality. Přesto je ale vhodné uvést alespoň ve stručnosti přibližnou dobu nutnou pro zpracování snímků. Měření rychlosti probíhalo na notebooku MacBook Pro 2017 s dvoujádrovým procesorem Intel Core i5[2].

rozlišení obrazu	1280x720		1920x1080	
počet snímků	20 snímků	200 snímků	20 snímků	200 snímků
doba trvání	2,94 s	27,06 s	6,82 s	64,89 s

Tabulka 6.2: Doba zpracování různě dlouhých sekvencí. Získané hodnoty jsou vždy aritmetickým průměrem deseti měření.

Jelikož implementovaný program využívá pro výpočty temporálních statistik jednoduchý filtru typu IIR prvního řádu, závisí celková rychlost mapování tónů pouze na rozlišení obrazu a počtu vstupních snímků. Hodnota konfigurovatelných parametrů rychlost výpočtu neovlivňuje. Samotná obrazová data mohou algoritmus zpomalit pouze v případě, že obsahují pixely se zápornými, či nekonečně velkými hodnotami. Kritickou operací z hlediska doby provádění mapování tonality je výpočet hodnot na základě percentilu, jelikož je nutné nejprve všechny obrazové body seřadit podle jejich velikosti.

Protože zpracování snímků implementovaný algoritmus rozděluje na dílčí kroky, je možné změřit čas potřebný k jejich vykonání. Následující tabulka ukazuje procentuální podíl doby výpočtu jednotlivých kroků vůči celkové době zpracování snímku:

fáze mapování tónů	**podíl z celkové doby**
předzpracování	10,8 %
dekompozice obrazu	39,8 %
zpracování detailů	5,3 %
komprese tonality	28,6 %
obnovení barev	15,5 %

Tabulka 6.3: Doba trvání dílčích kroků zpracování snímků. Měření probíhalo na sekvenci *showgirl_02* [32] o 341 snímcích s rozlišením 1920x1080 pixelů. Přibližně pětina snímků z této sekvence obsahuje záporné hodnoty obrazových bodů a vyžaduje výpočetně náročnější předzpracování. Uvedené podíly byly opět získány jako průměry deseti měření.

V tabulce 6.3 lze pozorovat, že nejdéle trvajícím krokem celého výpočtu je dekompozice obrazu. Obraz je dekomponován pomocí řízeného filtru, který je na obraz postupně aplikován dvakrát. Druhou nejdelší fází je komprese tonality, jelikož zde probíhá výpočet čtyř hodnot pomocí percentilu, kdy je nutné seřadit všechny pixely obrazu podle jejich velikosti. I přestože byla metoda pro výpočet percentilu zásadním způsobem optimalizována (viz podkapitolu 5.4.1) a řazení bodů obrazu proběhne ve fázi komprese tónů pouze dvakrát, stále se jedná o časově náročnou činnost. Důkladné zkoumání odhalilo, že výpočty hodnot na bázi percentilu tvoří přibližně polovinu délky trvání kroku komprese tonality.

[2]Technická specifikace procesoru Intel Core i5-7360U – https://ark.intel.com/content/www/us/en/ark/products/97535/intel-core-i5-7360u-processor-4m-cache-up-to-3-60-ghz.html.

Kapitola 7

Závěr

Hlavním cílem této diplomové práce bylo implementovat temporálně pracující operátor pro mapování tónů videa s vysokým dynamickým rozsahem a uvést teoretické informace související s touto problematikou. Aby bylo možné dosáhnout kvalitních výsledků, bylo nutné prostudovat rozsáhlé množství odborné literatury, která velmi často přesahovala rámec oboru zpracování obrazu. Všechny cíle zadání byly naplněny v plném rozsahu, pro aplikaci vzniklo i přehledné grafické rozhraní, které umožňuje přímočaré zkoumání vlivu konfigurovatelných parametrů na výsledný obraz.

Implementovaný program dokáže načíst libovolnou sekvenci snímků s vysokým dynamickým rozsahem ve formátu OpenEXR či Radiance HDR, na kterou aplikuje temporální mapování tónů zohledňující uživatelem definované parametry. Výstupem aplikace je při vhodné konfiguraci parametrů video neobsahující nežádoucí obrazové artefakty, které lze zobrazit na běžném zařízení jako je například počítačový monitor nebo konvenční televizor. Operátor, ze kterého implementace vychází, byl rozšířen o dodatečnou temporální stabilitu a jeho rozhraní bylo zjednodušeno pro zvýšení pohodlí uživatele při ovládání aplikace. Program dokáže oproti původnímu návrhu operátoru pracovat s libovolně uloženými vstupními daty, jež respektují podporované formáty, a nevyžaduje provedení externí normalizace snímků před jeho použitím. Podporovány jsou i snímky se zápornými a nekonečně velkými hodnotami pixelů. Byla také přidána možnost lokálního mapování tónů statických obrázků.

Samotná výstupní videa disponují při adekvátním nastavení parametrů velmi dobrými kontrastními vlastnostmi a neobsahují viditelné temporální artefakty. Rychlost temporální adaptace, která udává, jak rychle dokáže operátor reagovat na změny osvětlení scény napříč vstupními snímky, lze konfigurovat pomocí jednoho dedikovaného parametru. Celý algoritmus včetně předpracování snímků byl navržen tak, aby bylo možné jej případně aplikovat v reálném čase. Vliv uživatelských parametrů byl podrobně demonstrován a výsledky byly vizuálně porovnány s nejvyspělejšími existujícími operátory temporálního mapování tónů.

Program byl koncepčně navržen tak, aby bylo možné snadno rozšířit jeho funkcionalitu o další operátory mapování tónů. Modulární struktura, objektově orientovaný návrh a bohatě komentovaný zdrojový kód umožňují velmi snadné přidávání nových funkcí. Pro zrychlení výpočtu by bylo možné program akcelerovat na hardwaru, či dodatečně paralelizovat některé sekvenční kroky výpočtu. Jestliže by bylo cílem uživateli poskytnout maximální kontrolu nad algoritmem operátoru, mohl by být dodatečně implementován „pokročilý režim“ ovládání, který by v grafickém rozhraní zpřístupnil další vybrané parametry za účelem možnosti jejich konfigurace.

Literatura

[1] Alliance for Open Media: *Alliance for Open Media Established to Deliver Next-Generation Open Media Formats*. [online; cit. 3. 11. 2018]. Dostupné z https://aomedia.org/alliance-to-deliver-next-generation-open-media-formats/.

[2] Alliance for Open Media: *Members*. [online; cit. 3. 11. 2018]. Dostupné z https://aomedia.org/membership/members/.

[3] Apple Inc.: *Use HDR on your iPhone, iPad, and iPod touch*. [online; cit. 12. 11. 2018]. Dostupné z https://support.apple.com/en-us/HT207470.

[4] Aydin, T. O.; Nikolce Stefanoski, S.; Croci; aj.: Temporally Coherent Local Tone Mapping of HDR Video. *ACM Transactions on Graphics*, ročník 33, č. 6, listopad 2014: s. 196:1–196:13, ISSN 0730-0301, doi:10.1145/2661229.2661268.

[5] Banterle, F.; Corsini, M.; Cignoni, P.; aj.: A Low-Memory, Straightforward and Fast Bilateral Filter Through Subsampling in Spatial Domain. *Computer Graphics Forum*, ročník 31, č. 1, únor 2012: s. 19–32, ISSN 0167-7055, doi:10.1111/j.1467-8659.2011.02078.x.

[6] Bennett, E. P.; McMillan, L.: Video Enhancement Using Per-pixel Virtual Exposures. *ACM Transactions on Graphics*, ročník 24, č. 3, červenec 2005: s. 845–852, ISSN 0730-0301, doi:10.1145/1073204.1073272.

[7] Boitard, R.; Bouatouch, K.; Cozot, R.; aj.: Temporal Coherency for Video Tone Mapping. In *Proc. SPIE 8499, Applications of Digital Image Processing XXXV*, ročník 8499, říjen 2012, ISBN 9780819492166, s. 84990D:1–84990D:10, doi:10.1117/12.929600.

[8] Boitard, R.; Cozot, R.; Thoreau, D.; aj.: Survey of Temporal Brightness Artifacts in Video Tone Mapping. *HDRi2014 - Second International Conference and SME Workshop on HDR imaging (2014)*, leden 2014.

[9] Boitard, R.; Cozot, R.; Thoreau, D.; aj.: Zonal Brightness Coherency for Video Tone Mapping. *Image Communication*, ročník 29, č. 2, únor 2014: s. 229–246, ISSN 0923-5965, doi:10.1016/j.image.2013.10.001.

[10] Bonneel, N.; Tompkin, J.; Sunkavalli, K.; aj.: Blind Video Temporal Consistency. *ACM Transactions on Graphics*, ročník 34, č. 6, listopad 2015: s. 196:1–196:9, ISSN 0730-0301, doi:10.1145/2816795.2818107.

[11] Borer, T.; Cotton, A.: *A „Display Independent“ High Dynamic Range Television System*. [online; cit. 31. 10. 2018].
Dostupné z http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP309.pdf.

[12] Braun, R.; Garlington, S.: *High Dynamic Range Photos with Your DJI Drone*. [online; cit. 29. 10. 2018].
Dostupné z https://djiphotoacademy.com/hdr-photos-with-your-dji-drone/.

[13] Čadík, M.; Aydin, T. O.; Myszkowski, K.; aj.: Image Attributes and Quality for Evaluation of Tone Mapping Operators. In *Proceedings of Pacific Graphics 2006 (14th Pacific Conference on Computer Graphics and Applications)*, říjen 2006, s. 35–44.

[14] Čadík, M.; Aydin, T. O.; Myszkowski, K.; aj.: On evaluation of video quality metrics: an HDR dataset for computer graphics applications. In *Human Vision and Electronic Imaging XVI (Proceedings Volume)*, ročník 7865, editace B. E. Rogowitz; T. N. Pappas, SPIE, únor 2011, ISBN 978-0-8194-8402-4, s. 1–9, doi:10.1117/12.878875.

[15] Čadík, M.; Herzog, R.; Mantiuk, R.; aj.: New Measurements Reveal Weaknesses of Image Quality Metrics in Evaluating Graphics Artifacts. *ACM Transactions on Graphics*, ročník 31, č. 6, listopad 2012: s. 147:1–147:10, ISSN 0730-0301, doi:10.1145/2366145.2366166.

[16] Debevec, P.; Malik, J.: Recovering High Dynamic Range Radiance Maps from Photographs. *ACM Transactions on Graphics*, srpen 1997: s. 369–378, ISSN 0730-0301, doi:10.1145/258734.258884.

[17] Dehaene, S.: The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in cognitive sciences*, ročník 7, květen 2003: s. 145–147, ISSN 1364-6613, doi:10.1016/S1364-6613(03)00055-X.

[18] Dolby Laboratories: *An Introduction to Dolby Vision^{TM}*. [online; cit. 1. 11. 2018].
Dostupné z https://dolby.com/us/en/technologies/dolby-vision/an-introduction-to-dolby-vision.pdf.

[19] Dolby Laboratories: *Dolby Vision^{TM} for the Home*. [online; cit. 1. 11. 2018].
Dostupné z https://dolby.com/us/en/technologies/dolby-vision/dolby-vision-white-paper.pdf.

[20] Dong, X.; Bonev, B.; Zhu, Y.; aj.: Region-based temporally consistent video post-processing. In *The IEEE Conference on Computer Vision and Pattern Recognition*, červen 2015, ISBN 978-1-4673-6964-0, s. 714–722, doi:10.1109/CVPR.2015.7298671.

[21] Drago, F.; Myszkowski, K.; Annen, T.; aj.: Adaptive Logarithmic Mapping For Displaying High Contrast Scenes. *Computer Graphics Forum*, ročník 22, č. 3, září 2003: s. 419–426, ISSN 0167-7055, doi:10.1111/1467-8659.00689.

[22] Durand, F.; Dorsey, J.: Interactive Tone Mapping. In *Rendering Techniques 2000: Proceedings of the Eurographics Workshop in Brno, Czech Republic*, Springer, červen 2000, ISBN 978-3-7091-6303-0, s. 219–230, doi:10.1007/978-3-7091-6303-0_20.

[23] Durand, F.; Dorsey, J.: Fast Bilateral Filtering for the Display of High-dynamic-range Images. *ACM Transactions on Graphics*, ročník 21, č. 3, červenec 2002: s. 257–266, ISSN 0730-0301, doi:10.1145/566654.566574.

[24] Eilertsen, G.; Mantiuk, R. K.; Unger, J.: A comparative review of tone-mapping algorithms for high dynamic range video. *Computer Graphics Forum*, ročník 36, č. 2, květen 2017: s. 565–592, ISSN 1467-8659, doi:10.1111/cgf.13148.

[25] Eilertsen, G.; andJonas Unger, R. M.: Real-time noise-aware tone mapping. *ACM Transactions on Graphics*, ročník 34, říjen 2015: s. 1–15, ISSN 0730-0301, doi:10.1145/2816795.2818092.

[26] Eilertsen, G.; Wanat, R.; Mantiuk, R. K.; aj.: Evaluation of Tone Mapping Operators for HDR-Video. *Computer Graphics Forum*, ročník 32, č. 7, 2013: s. 275–284, ISSN 0167-7055, doi:10.1111/cgf.12235.

[27] El Mezeni, D.; Saranovac, L.: Enhanced local tone mapping for detail preserving reproduction of high dynamic range images. *Journal of Visual Communication and Image Representation*, ročník 53, březen 2018: s. 122–133, ISSN 1047-3203, doi:10.1016/j.jvcir.2018.03.007.

[28] El Mezeni, D.; Saranovac, L.: Temporal adaptation control for local tone mapping operator. *Journal of Electrical Engineering*, ročník 69, srpen 2018: s. 261–269, ISSN 1339-309X, doi:10.2478/jee-2018-0037.

[29] Fattal, R.; Lischinski, D.; Werman, M.: Gradient Domain High Dynamic Range Compression. *ACM Transactions on Graphics*, ročník 21, 07 2002: s. 249–256, ISSN 0730-0301, doi:10.1145/566570.566573.

[30] Ferwerda, J.; Pattanaik, S. N.; Shirley, P.; aj.: A Model of Visual Adaptation for Realistic Image Synthesis. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, ročník 30, červenec 1996, ISBN 0-89791-746-4, s. 249–258, doi:10.1145/237170.237262.

[31] Feyiz, K.; Kamisli, F.; Zerman, E.; aj.: Statistical analysis and directional coding of layer-based HDR image coding residue. In *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*, 10 2017, ISSN 2473-3628, s. 1–6, doi:10.1109/MMSP.2017.8122244.

[32] Froehlich, J.; Grandinetti, S.; Eberhardt, B.; aj.: Creating Cinematic Wide Gamut HDR-Video for the Evaluation of Tone Mapping Operators and HDR-Displays. In *Proc. SPIE 9023, Digital Photography X*, ročník 9023, únor 2014, s. 9023:1–9023:10, doi:10.1117/12.2040003.

[33] Fulton, W.: *EV - Exposure Value (photography)*. [online; cit. 28. 10. 2018]. Dostupné z https://scantips.com/lights/evchart.html.

[34] Ganong, W. F.: *Přehled lékařské fyziologie*. Galén, 2005, ISBN 80-7262-311-7, 156–176 s.

[35] Gilchrist, A.; Jacobsen, A.: Perception of Lightness and Illumination in a World of One Reflectance. *Perception*, ročník 13, č. 1, 1984: s. 5–19, ISSN 0301-0066, doi:10.1068/p130005.

[36] He, K.; Sun, J.; Tang, X.: Guided Image Filtering. *IEEE transactions on pattern analysis and machine intelligence*, ročník 35, červen 2013: s. 1397–1409, ISSN 0162-8828, doi:10.1109/TPAMI.2012.213.

[37] Hirsch, R.: *Light and Lens: Photography in the Digital Age*. Taylor & Francis, 2018, ISBN 9781317371700.

[38] International Telecommunication Union: *BT.2020: Parameter values for ultra-high definition television systems for production and international programme exchange.* [online; cit. 31. 10. 2018].
Dostupné z https://itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2020-2-201510-I!!PDF-E.pdf.

[39] International Telecommunication Union: *BT.2100-2: Image parameter values for high dynamic range television for use in productionand international programme exchange.* [online; cit. 31. 10. 2018].
Dostupné z https://itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2100-2-201807-I!!PDF-E.pdf.

[40] International Telecommunication Union: *BT.709-6: Parameter values for the HDTV standards for production and international programme exchange.* [online; cit. 31. 10. 2018].
Dostupné z https://itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.709-6-201506-I!!PDF-E.pdf.

[41] Irawan, P.; Ferwerda, J. A.; Marschner, S.: Perceptually Based Tone Mapping of High Dynamic Range Image Streams. In *Proceedings of the Sixteenth Eurographics Conference on Rendering Techniques*, prosinec 2005, ISBN 3-905673-23-1, s. 231–242, doi:10.2312/EGWR/EGSR05/231-242.

[42] Kainz, F.; Bogart, R.; Stanczyk, P.: *Technical Introduction to OpenEXR.* [online; cit. 31. 10. 2018].
Dostupné z http://openexr.com/documentation/TechnicalIntroduction.pdf.

[43] Kang, H.; Lee, S. H.; Song, K. S.; aj.: Bayer patterned high dynamic range image reconstruction using adaptive weighting function. *EURASIP Journal on Advances in Signal Processing*, ročník 12, č. 2, listopad 2014: s. 76:1–76:18, ISSN 1687-6180, doi:10.1186/1687-6180-2014-76.

[44] Kuang, J.; Johnson, G. M.; Fairchild, M. D.: iCAM06: A refined image appearance model for HDR image rendering. *Visual Communication and Image Representation*, ročník 18, říjen 2007: s. 406–414, ISSN 1047-3203.

[45] Kumbhar, K.; Kshirsagar, K. P.: Comparative Study of CCD & CMOS Sensors for Image Processing. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, ročník 3, č. 12, prosinec 2015: s. 194–196, ISSN 2321-5526.

[46] Lang, M.; Wang, O.; Aydin, T.; aj.: Practical Temporal Consistency for Image-based Graphics Applications. *ACM Transactions on Graphics*, ročník 31, č. 4, červenec 2012: s. 34:1–34:8, ISSN 0730-0301, doi:10.1145/2185520.2185530.

[47] Ledda, P.; Santos, L. P.; Chalmers, A.: A Local Model of Eye Adaptation for High Dynamic Range Images. In *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, AFRIGRAPH '04, 2004, ISBN 1-58113-863-6, s. 151–160, doi:10.1145/1029949.1029978.

[48] Lee, C.; Kim, C.-S.: Gradient Domain Tone Mapping of High Dynamic Range Videos. In *Proceedings of International Conference on Image Processing*, ročník 3, leden 2007, ISBN 978-1-4244-1437-6, s. III–461–III–464, doi:10.1109/ICIP.2007.4379346.

[49] Liu, Y.; Wang, C.; Zhao, H.; aj.: Bayer Image Demosaicking Using Eight-Directional Weights Based on the Gradient of Color Difference. *Symmetry*, ročník 10, červen 2018: s. 222–237, ISSN 2073-8994, doi:10.3390/sym10060222.

[50] Loucký, M.: *HDR10+ je tu pro nejlepší divácký zážitek v UHD rozlišení*. [online; cit. 1. 11. 2018].
Dostupné z https://chip.cz/temata/hdr10-je-tu-pro-nejlepsi-divacky-zazitek-v-uhd-rozliseni/.

[51] Macenauer, A.: *CMOS čidla a jeho vlastnosti*. [online; cit. 27. 10. 2018].
Dostupné z https://fotoaparat.cz/clanek/234/cmos-cidla-a-jejich-vlastnosti-5022/.

[52] Macenauer, A.: *Jak funguje CCD čidlo*. [online; cit. 27. 10. 2018].
Dostupné z https://fotoaparat.cz/clanek/233/jak-funguje-ccd-cidlo-5021/.

[53] Mai, Z.; Mansour, H.; Mantiuk, R.; aj.: Optimizing a Tone Curve for Backward-Compatible High Dynamic Range Image and Video Compression. *IEEE Transactions on Image Processing*, ročník 20, č. 6, červen 2011: s. 1558–1571, ISSN 1057-7149.

[54] Mantiuk, R.; Daly, S.; Kerofsky, L.: Display Adaptive Tone Mapping. *ACM Transactions on Graphics*, ročník 27, č. 3, srpen 2008: s. 68:1–68:10, ISSN 0730-0301, doi:10.1145/1360612.1360667.

[55] Mantiuk, R.; Krawczyk, G.; Myszkowski, K.; aj.: Perception-motivated high dynamic range video encoding. *ACM Transactions on Graphics*, ročník 23, č. 3, květen 2004: s. 733–741, ISSN 0730-0301, doi:10.1145/1015706.1015794.

[56] McConnell, J.: *An Invisible Tax on the Web: Video Codecs*. [online; cit. 3. 11. 2018].
Dostupné z https://blog.mozilla.org/blog/2018/07/11/royalty-free-web-video-codecs/.

[57] MulticoreWare Inc.: *About x265*. [online; cit. 3. 11. 2018].
Dostupné z http://x265.org/about/.

[58] Nayar, S. K.; Mitsunaga, T.: High dynamic range imaging: spatially varying pixel exposures. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, ročník 1, červen 2000, ISBN 0-7695-0662-3, s. 472–479, doi:10.1109/CVPR.2000.855857.

[59] Olšan, J.: *YouTube a Netflix už nabízejí videa ve formátu AV1, můžete si je vyzkoušet.* [online; cit. 3. 11. 2018].
Dostupné z https://cnews.cz/video-kodek-av1-beta-testovani-youtube-netflix.

[60] Ozer, J.: *Netflix Finds x265 20% More Efficient than VP9.* [online; cit. 3. 11. 2018].
Dostupné z http://streamingmedia.com/Articles/Editorial/Featured-Articles/Netflix-Finds-x265-20-More-Efficient-than-VP9-113346.aspx.

[61] Pihan, R.: *Oko (a mozek) versus fotoaparát.* [online; cit. 10. 10. 2018].
Dostupné z http://fotoroman.cz/tech1/light_eye_camera.htm.

[62] Rayburn, D.: *Google Has A Problem: VP8 Is Not As Good As H.264.* [online; cit. 3. 11. 2018].
Dostupné z https://businessinsider.com/google-has-a-problem-vp8-is-not-as-good-as-h264-2010-5.

[63] Reichl, J.; Všetička, M.: *Světlo jako elektromagnetické vlnění.* [online; cit. 11. 10. 2018].
Dostupné z http://fyzika.jreichl.com/main.article/view/434-svetlo-jako-elektromagneticke-vlneni.

[64] Reinhard, E.; Stark, M.; Shirley, P.; aj.: Photographic Tone Reproduction for Digital Images. *ACM Transactions on Graphics*, ročník 21, č. 3, červenec 2002: s. 267–276, ISSN 0730-0301, doi:10.1145/566654.566575.

[65] Reinhard, E.; aj.: *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting.* Morgan Kaufmann, 2010, ISBN 9780125852630, 1–4 s.

[66] Robertson, M. A.; Borman, S.; Stevenson, R. L.: Estimation-Theoretic Approach to Dynamic Range Enhancement Using Multiple Exposures. *Journal of Electronic Imaging*, ročník 12, č. 2, duben 2003: s. 219–228, ISSN 1017-9909.

[67] Samsung: *Samsung and Amazon Prime Video First to Launch HDR10+ Content.* [online; cit. 1. 11. 2018].
Dostupné z https://news.samsung.com/global/samsung-and-amazon-prime-video-first-to-launch-hdr10-content.

[68] Savazzi, E.: *Digital Photography for Science.* Lulu.com, 2011, ISBN 9780557911332, 209–211 s.

[69] Schlick, C.: Quantization Techniques for Visualization of High Dynamic Range Pictures. In *Photorealistic Rendering Techniques*, Springer, březen 1995, ISBN 978-3-642-87825-1, s. 7–20, doi:10.1007/978-3-642-87825-1_2.

[70] Schöberl, M.; Belz, A.; Seilerr, J.; aj.: High Dynamic Range Video by Spatially Non-Regular Optical Filtering. In *IEEE International Conference on Image Processing (ICIP)*, září 2012, s. 2757–2760, doi:10.1109/ICIP.2012.6467470.

[71] Schöberl, M.; Keinert, J.; Ziegler, M.; aj.: Evaluation of a high dynamic range video camera with non-regular sensor. In *Proc. SPIE 8660, Digital Photography IX*, únor 2013, s. 86600M–86600M–12, doi:10.1117/12.2004258.

[72] Seeman, M.; Zemčík, P.; Juránek, R.; aj.: Fast bilateral filter for HDR imaging. *Journal of Visual Communication and Image Representation*, ročník 2012, č. 1, 2011, ISSN 1047-3203, doi:10.1016/j.jvcir.2011.07.012.

[73] Shahid, H.; Li; Fanaswala; aj.: A new hybrid tone mapping scheme for high dynamic range (HDR) videos. In *2015 IEEE International Conference on Consumer Electronics (ICCE)*, leden 2015, ISSN 2158-3994, s. 351–352, doi:10.1109/ICCE.2015.7066441.

[74] Shilov, A.: *Alliance for Open Media Releases Royalty-Free AV1 1.0 Codec Spec.* [online; cit. 3. 11. 2018].
Dostupné z https://anandtech.com/show/12601/alliance-for-open-media-releases-royaltyfree-av1-10-codec-spec.

[75] Szeliski, R.: *Computer Vision: Algorithms and Applications.* Springer, 2011, ISBN 978-1-84882-935-0, 118 s.

[76] Tocci, M. D.; Kiser, C.; Tocci, N.; aj.: A Versatile HDR Video Production System. *ACM Transactions on Graphics (Proceedings of SIGGRAPH '11)*, ročník 30, č. 4, červenec 2011: s. 41:1–41:10, ISSN 0730-0301, doi:10.1145/2010324.1964936.

[77] Tomasi, C.; Manduchi, R.: Bilateral Filtering for Gray and Color Images. In *Proceedings of the Sixth International Conference on Computer Vision*, IEEE Computer Society, leden 1998, ISBN 81-7319-221-9, s. 839–846, doi:10.1109/ICCV.1998.710815.

[78] Tomaszewska, A.; Mantiuk, R.: *Image Registration for Multi-exposure High Dynamic Range Image Acquisition.* Szczecin University of Technology, listopad 2007.

[79] Tumblink, J.; Turk, G.: LCIS: A Boundary Hierarchy for Detail-preserving Contrast Reduction. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '99, ACM, 1999, ISBN 0-201-48560-5, s. 83–90, doi:10.1145/311535.311544.

[80] Tursun, O. T.; Akyüz, A. O.; Erdem, A.; aj.: The State of the Art in HDR Deghosting: A Survey and Evaluation. *Comput. Graph. Forum*, ročník 34, č. 2, květen 2015: s. 683–707, ISSN 0167-7055, doi:10.1111/cgf.12593.

[81] Ward, G.: *High Dynamic Range Image Encodings.* [online; cit. 30. 10. 2018].
Dostupné z http://anyhere.com/gward/hdrenc/Encodings.pdf.

[82] Ward Larson, G.; Rushmeier, H. E.; Piatko, C.: A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, ročník 3, říjen 1997: s. 291–306, ISSN 1077-2626, doi:10.1109/2945.646233.

[83] Xiao, K.; Fu, C.; Sophie, D. K.; aj.: Visual gamma correction for LCD displays. *Displays*, ročník 32, 01 2011: s. 17–23, ISSN 0141-9382, doi:10.1016/j.displa.2010.09.003.

[84] Zhao, S.: *Wavelength of Maximum Human Visual Sensitivity.* [online; cit. 20. 10. 2018].
Dostupné z https://hypertextbook.com/facts/2007/SusanZhao.shtml.

Přílohy

Příloha A

Obsah přiloženého paměťového média

- `bin/` – spustitelná podoba aplikace pro operační systém MacOS
- `data_set/` – vybrané vstupní sekvence
- `doc/` – zdrojové soubory textové části práce
- `lib/` – zdrojové soubory knihovny OpenCV 3
- `README.md` – stručný manuál k překladu zdrojového kódu aplikace
- `result/` – videa demonstrující dosažitelné výsledky
- `src/` – zdrojové soubory implementované aplikace
- `xstane33_dip_print.pdf` – text diplomové práce určený zejména k tisku, obsahuje obrázky s vysokým rozlišením a bezbarvé odkazy
- `xstane33_dip_display.pdf` – text diplomové práce obsahující barevné odkazy a obrázky s nižším rozlišením