Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Zlepšenie kvality obrazov z bezpečnosných kamier

Bakalárska práca

2023
Ondrej Bublavý

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Zlepšenie kvality obrazov z bezpečnosných kamier

Bakalárska práca

Študijný program:	Aplikovaná informatika
Študijný odbor:	Informatika
Školiace pracovisko:	Katedra informatiky
Školiteľ:	RNDr. Zuzana Černeková, PhD

Bratislava, 2023
Ondrej Bublavý

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:
Študijný program:

Študijný odbor:
Typ záverečnej práce:
Jazyk záverečnej práce:
Sekundárny jazyk:

Názov:

Anotácia:

Vedúci:
Katedra:
Vedúci katedry:

Dátum zadania:

Dátum schválenia:

garant študijného programu

.....

študent

.....

vedúci práce

Poďakovanie: Tu môžete poďakovať školiteľovi, prípadne ďalším osobám, ktoré vám s prácou nejako pomohli, poradili, poskytli dáta a podobne.

Abstrakt

Problematika kvality obrazov z bezpečnostných kamier má množstvo špecifík, ktoré ju vynímajú z klasickej témy spracovania digitálneho obrazu a videa. Podlieha charakteristickej množine scenárov - zaberá predmetnú scénu vždy zo vzdialenosti rádovo v metroch až desiatkach metrov, musí čeliť svetelným a často aj poveternostným podmienkam nepretržite s plynutím každej hodiny, dňa a mesiaca, a predovšetkým neslúži k poskytnutiu esteticky pútavého, čo najviac realistického či kinematického vizuálu svojej scény, ale k identifikácii osôb, vozidiel a aktivít na nej situovaných. Táto problematika teda čelí špecifickej sade výziev, a teda si vyžaduje osobitné prístupy a našli v nej uplatnenie inde nevyužívané algoritmy. Zároveň aj tu priniesli osoh mnohé z metód, zamýšľaných k všeobecnému zvyšovaniu kvality niektorých z atribútov videa. V tejto práci sme preskúmali paletu relevantných prístupov a algoritmov, implementovali tie, ktoré majú zverejnené verzie a kódy nasaditeľné do praxe, a to v rámci jedného programu, ktorý umožňuje tieto metódy kombinovať - aj s nastavením ich parametrov - a otestovali postupnosti ich aplikácií poťažmo k bežným scenárom v akých bývajú bezpečnostné kamery nasadené, a testovacím videám v týchto scenároch zhotoveným. Našli sme súvislosti určujúce vhodnosť použitia jednotlivých algoritmov a ich kombinácií na základe prostredia a inherentných vlastností videa z daného typu bezpečnostnej kamery. Dosiahli sme viditeľné zlepšenie v čitateľnosti čŕt a podstatných detailov objektov záujmu a demonštrovali sme tieto výsledky porovnaniami, z ktorých ukážky sú zahrnuté v práci. Okrem obrazových vlastností sme posudzovali aj časovú náročnosť metód, a teda ich vhodnosť použitia na zlepšenie priameho prenosu, alebo ich aplikovateľnosť len na už nahraný záznam.

Kľúčové slová: bezpečnostné kamery, dohľadové systémy, kvalita videa, zvyšovanie rozlíšenia, zvyšovanie kontrastu, idenitifkácia osôb

Abstract

The issue of the quality of images from security cameras has a number of specifics that distinguish it from the classical topic of digital image and video processing. It is subject to a characteristic set of scenarios - it monitors the surveilled scene from a distance on the order of meters to tens of meters, it has to face light and often weather conditions continuously with the passage of each hour, day and month, and above all, it's goal is not to provide an aesthetically appealing, most realistic, or cinematic visual of its scene, but to identify the people, vehicles and activities situated on it. This area thus faces a specific set of challenges, and therefore requires specific approaches, and algorithms not used elsewhere have found application here. At the same time, many of general methods intended to enhance the quality of some of the common video attributes have also been beneficial here. In this work, we have explored a variety of relevant approaches and algorithms, implemented those released and deployable for practical usage, within a singular program that allows these methods to be combined - including setting up their parameters - and we tested the sequences and orders of their application in common scenarios, in which security cameras tend to be deployed, and on test videos taken in these scenarios. We found characteristics determining the eligibility of using these algorithms and their combinations based on the environment and inherent attributes of the video from the given type of security camera. We have achieved a noticeable improvement in the recognizability of features and essential details of the objects of interest, and have demonstrated these results with comparisons, examples of which are included in the paper. In addition to the visual properties, we have also assessed time demanded by the methods, and hence their suitability for deployment with footage streamed in real time, or their applicability only to already captured recording.

Keywords: security cameras, surveillance systems, video enhancement, super-resolution, contrast enhancement, person identification

Obsah

Zoznam obrázkov

Zoznam tabuliek

Úvod

Boj s kriminalitou a bezpečnosť verejných, pracovných či súkromných priestorov je od nepamäti implicitnou témou náležiacou ľudskej civilizácii. Dalo by sa polemizovať, že bezpečnostné kamery v nej priniesli najväčší prielom od čias kodifikácie prvých zákonov, keďže umožnili nesporné dokazovanie ich porušenia alebo neporušenia, 24 hodín denne, 7 dní v týždni. Kvalita obrazu z bezpečnostných kamier je pritom dôležitá z perspektívy prítomnosti, minulosti, ako aj budúcnosti. Dostatočne detailný záznam v reálnom čase pracovníkovi poverenému dohľadom nad priestorom dovoľuje identifikovať pohyb osôb, vozidiel, iných objektov, a výkon aktivít v ňom, bez nutnosti jeho fyzickej prítomnosti, a teda aj na viacerých miestach či z viacerych uhlov pohľadu súčasne. Rovnako môže slúžiť automatizovanému klasifikačnému, prípadne poplašnému systému, využívajúcemu kamerové systémy pre sledovanie tohto priestoru. Z pohľadu minulosti umožňuje spoľahlivé vyšetrenie nehodových udalostí či bezpečnostných incidentov. A do budúcnosti im tak umožňuje predchádzať – či už aktívne – napravením predošlých chýb v ochrane priestoru, alebo pasívne – odradením páchateľov.

Doposiaľ k dátumu zadania tejto práce nebol navrhnutý algoritmus, ani definovaná metodológia, vedúca k všeobecnému zlepšeniu kvality výstupných obrazov z bezpečnostných kamier. Prezentované algoritmy relevantné v tejto problematike sa buď zapodievali len jednou z možných oblastí, ktorými určujeme kvalitu videa a identifikovateľnosť objektov a aktivít na ňom zachytených, nehľadiac na kombinovateľnosť s prístupmi cieliacimi na iné zlepšiteľné atribúty, alebo sa ani nezaoberali konkrétne doménou bezpečnostných kamier a jej špecifikami, ale digitálnym videom ako takým. My sa pokúsime zlepšenie kvality záznamov z dohľadových systémov poňať ako samonosnú úlohu s jednotným cieľom, ktorý posudzujeme zlepšením schopnosti ľudského alebo programového pozorovateľa identifikovať objekty a kontext v možných scenároch v rámci nasadenia bezpečnostných kamier.

V kapitole 1 identifikujeme výzvy, problémy a atribúty, v ktorých zlepšenie prispieva k zlepšeniu celkovej kvality záznamu. Taktiež sa v nej pozrieme na už existujúce prístupy a metódy a na ich relevanciu pre našu oblasť. V kapitole 2 uvedieme, aké algoritmy implementujeme pre účely tejto práce, a akým spôsobom a na akých videách ich testovať. Porovnávať pritom nebudeme - narozdiel od mnohých prác na tému spracovania obrazu - iba samostatné algoritmy, ale najmä rôzne kombinácie týchto al-

goritmov, vrátane poradí ich nasadení, v zmysle vytýčeného cieľa dosiahnuť zlepšenie presahujúce viacero kvalitatívnych kritérií. Výsledky týchto porovnaní na testovacích videách spadajúcich do širšej palety realistických a bežných scenárov budeme prezentovať v kapitole 3, kde zhodnotíme aj ich nameranú časovú náročnosť, a teda možnosť nasadenia na živý prenos v reálnom čase a nie len na už zhotovený záznam.

Kapitola 1

Výzvy a existujúce prístupy

Aby bol záznam z bezpečnostnej kamery čo najnápomocnejší pri predchádzaní, odhaľovaní a vyšetrovaní škodnej či priamo kriminálnej činnosti, musí byť pri jeho vyhotovovaní, prípadnom ukladaní a zobrazovaní, kladený dôraz na určité atribúty jeho kvality, či už má záznam slúžiť ľudskému pozorovateľovi, alebo nejakému rozpoznávaciemu algoritmu. Aktuálne neexistuje žiadna norma ani štandard definujúci tieto atribúty, avšak v rozsahu našej práce si ich definujeme, aby sme si vedeli orientačne stanoviť metriky, zlepšenie v ktorých ústi zároveň k zlepšeniu v praktickom nasadení. Pozrime sa teda na aspekty, ktoré budeme sledovať.

1.1 Atribúty kvality záznamu

Kontext kvality obrazu systémov bezpečnostných kamier je podmnožinou oblasti kvality videozáznamu, ktorý je podmnožinou oblasti kvality digitálneho obrazu. Preto niektoré atribúty, ktorými sa budeme zaoberať, sú zároveň predmetom záujmu zmienených nadmnožín, a niektoré naberajú osobitý dôraz v odvetví bezpečnostných kamier.

1.1.1 Všeobecne relevantné atribúty kvality videa

- Jednou zo základných metrík kvality digitálneho obrazu je rozlíšenie. V dnešnej dobe už sú dostupné senzory vysokých rozlíšení za relatívne nízke ceny, za čo vďačíme aj takzvanej honbe za megapixelmi. Samotné rozlíšenie je pri bezpečnostných kamerách výzvou v dvoch hlavných prípadoch. Jeden z prípadov, kedy potrebujeme vyššie rozlíšenie, než je priemyselným štandardom (najčastejšie 720p a 1080p), je, keď je kamera umiestnená tak, aby pozorovala rozsiahly priestor (napríklad veľké parkovisko), a teda veľkosť detailov, ako ľudské tváre či evidenčné čísla vozidiel, je v pomere k celému záberu nepostačujúca. Algoritmy spoľahlivo dopočítavajúce pixely obrazu, škálujúc ho na vyššie rozlíšenie tak, aby došlo k zlepšeniu rozoznateľnosti detailu a nie len k pocitu, že obraz ako celok

je kvalitnejší, no neobsahuje viac informácii, sú stále v podstate v plienkach. Vzostup umelej inteligencie a nárast miery jej chápania kontextu vizuálnej informácie to však rýchlo mení. Jedným z prípadov, kedy si zas vysoké rozlíšenie pri zázname nemôžeme dovoliť je, keď je záznam z bezpečnostnej kamery ukladaný na pamäťové médium. To má totiž obvykle obmedzenú kapacitu a zvýšenie rozlíšenia ukladaného obrazu bude mať za následok zníženie dĺžky záznamu, ktorú dokážeme na pamäťovom médiu uchovať. Zaujímavým prístupom v tejto oblasti je udalosťami riadená adaptácia, kedy sa systém snaží rozlišovať časové úseky s nízkou či neexistujúcou aktivitou (zmenou na zábere), ktoré nie je nutné ukladať vo vysokej kvalite, a časové úseky, kedy potenciálne dochádza k udalosti – zmene pozorovaného stavu – ktorú môžeme mať záujem zachovať v zázname vysokej kvality, pre prípad budúceho prešetrenia [6].

- Okrem rozlíšenia hrá v rovnici kvality záznamu a dostupného úložiska rolu aj ďalšia premenná – počet snímok za sekundu. Aj na tento parameter vieme aplikovať udalosťami riadenú adaptáciu. Okrem toho sú na vzostupe algoritmy dopočítavajúce snímky medzi dvojicami skutočne zaznamenaných záberov, tu však opäť platí, že je veľmi obtiažne a v praxi obvykle nemožné dopočítať informáciu tam, kde nám na originálnom zázname chýba. Tieto algoritmy len zvyšujú vnímanú plynulosť záznamu, no nedokážu pomôcť odhaliť skutočnosti, ktoré sa na pôvodnom zázname neobjavili.

- Bitová hĺbka nám hovorí o rozsahu farieb a ich jasov, ktorými dokážeme reprezentovať obraz v digitálnej podobe. Väčší význam v aplikácii bezpečnostných kamier však má dynamický rozsah. Ten nám hovorí o schopnosti samotnej kamery a jej spracúvajúceho softvéru zachytiť čo najširšie spektrum jasností. Dynamický rozsah v reálnych prostrediach presahuje rozsah reprezentovaný v 8-bitových textúrnych mapách [7][8]. Je veľmi dôležitý pri práci s prirodzeným osvetlením, kedy jasnosť prostredia dosahuje od rádovo niekoľkých nitov po rádovo desaťtisíce nitov [9].

- Pomer signálu a šumu je významný pre subjektívne vnímanú kvalitu obrazu. Značne zašumený obraz komplikuje ľudskému pozorovateľovi, a o to viac prípadne nasadeným rozpoznávacím algoritmom, schopnosť identifikovať objekty, ich hranice a črty, no aj ich pohyb. Metódy na odstraňovanie šumu pomáhajú zvýrazneniu skutočnej informácie, nedotvárajú však stratenú informáciu.

- Ostrosť je veľmi dôležitá metrika v kontexte bezpečnostných kamier, keďže určuje výraznosť (kontrast) hrán v obraze, a teda aj separovateľnosť a identifikovateľnosť čŕt a objektov.

Obr. 1.1: Porovnanie záberu zachytávajúceho scénu vo vysokom (naľavo) a v nízkom (napravo) dynamickom rozsahu [1]

1.1.2 Atribúty kvality osobité bezpečnostným kamerám

- Zorný uhol v praxi hovorí o tom, aký rozsiahly a prípadne zakrivený priestor dokážeme pokryť jednou kamerou. Táto metrika je však určená dostupným hardvérom, od softvéru závisí až následné spracovanie zakrivenia, prípadne rozmazania a skreslenia na krajoch obrazu.

- Vlastnosti pri nízkej hladine osvetlenia naberajú obzvlášť veľký význam pri bezpečnostných kamerách, nakoľko značná časť záznamu bude pri väčšine aplikácii vyhotovovaná za tmy. Ako veľmi záleží na schopnostiach kamery a obraz spracúvajúceho softvéru pri nízkom osvetlení závisí od prostredia, v ktorom je kamera použitá. Kamery, ktoré sú umiestnené v konštantne osvetlených priestoroch, najčastejšie interiéroch, alebo priestoroch kde je osvetlenie automaticky aktivované senzorom pohybu, môžu zanedbať aspekt použiteľnosti pri najnižších hladinách ambientného jasu. Dôraz na schopnosti v tme nemusia klásť ani kamery inštalované za účelom sledovania priestoru v čase, kedy sa v ňom očakáva zvýšená aktivita ľudí, ak je to len za denného svetla. Väčšina aplikácii je však spätá s nepretržitým záznamom a ak nie je kamerovému systému nápomocné automatické osvetľovanie priestoru v prípade aktivity detekovanej iným senzorom, sú schopnosti sprostredkovať črtovo a informačne čitateľný obraz za tmy minimálne rovnako dôležité, a často dôležitejšie, než vlastnosti obrazu cez deň – čo nie vždy platí pri posudzovaní kvality kamerového záznamu pre všeobecnejšie aplikácie.

Z tohto prehľadu nám vyplýva, že väčšina atribútov kvality je zhodná s všeobecnejšími aplikáciami kamerových záznamov, a teda mnohé algoritmy pôvodne nezamerané špecificky na bezpečnostné kamery, budú mať značný prínos aj v tejto podoblasti.

1.2 Existujúce prístupy a riešenia

Zhrnuli sme si hlavné oblasti, v ktorých vieme dosahovať zlepšenie pre celkovú kvalitu videozáznamu z bezpečnostných kamier. Predstavme si teraz myšlienky niektorých z najosožnejších prístupov venujúcich sa týmto oblastiam.

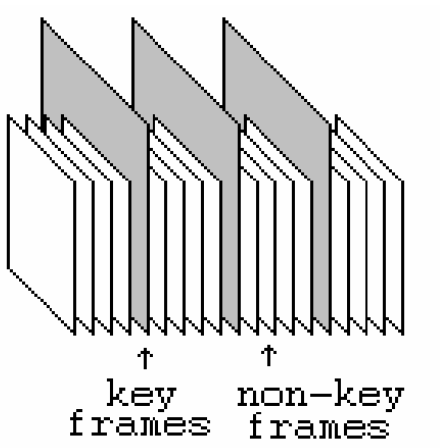
1.2.1 Zvýšenie rozlíšenia na základe databázy blokov vysokej kvality

Ako bolo zmienené v prvej časti kapitoly, môžeme mať k dispozícii kameru s vysokým rozlíšením, avšak nemať kapacitu ukladať snímky vo vysokom rozlíšení po dostatočne dlhú dobu. Predstavme si, že by sme chceli vyhotovovať 8K záznam s frekvenciou 30 snímok za sekundu a uchovávať ho po dobu 48 hodín, napríklad pre prípad vyšetrovania kriminálnej činnosti. Museli by sme ukladať 30 fps $\times$ 3600 sekúnd $\times$ 48 hodín = 4147200 snímok, pričom každá by bola o rozmere 7680 $\times$ 4320 pixelov, a na reprezentáciu každého pixelu by sme použili 24 bitov (8 bitov pre každý farebný kanál). Potrebovali by sme tak minimálne (4147200 $\times$ 7680 $\times$ 4320 $\times$ 24) / 8 bitov $\approx$ 28840 GB, teda skoro 30 TB. Toto číslo vieme niekoľkonásobne znížiť kompresiou videa, no stále budeme záznamom okupovať pamäť rádovo minimálne v pár terabajtoch. A čo ak by sme chceli záznam uchovávať nie dva dni, ale dva týždne[1]? Ako vieme výrazne ušetriť požiadavky na pamäť bez toho, aby sme sa vzdali množstva podstatných informácii vo videu? Jedným z možných riešení je uložiť všetky, alebo väčšinu snímok v zníženom rozlíšení. Na rekonštrukciu detailov v snímkach nižšieho rozlíšenia sa najčastejšie využíva databáza, v ktorej sa hľadá blok pixelov vo vysokom rozlíšení najpodobnejší danému zaostrovanému bloku po interpolácii. Na základe skladby databázy blokov vo vysokom rozlíšení, sú najčastejšie nasledujúce dva prístupy:

1.2.1.1 Ukladanie kľúčových snímok

Môžeme v originálnom rozlíšení ukladať každý k-ty (alebo iným vzorcom pravidelne zvolený) záber alebo statické snímky [10] a následne využiť tieto kľúčové zábery v pôvodnej kvalite k dopočítaniu nezachovaných detailov na ostatných snímkach. Tento koncept ilustruje obrázok 1.2 z článku [2]. Majme pri tom na vedomí, že všetky zábery, vrátane tých kľúčových, sa ukladajú komprimované. Avšak zábery, ktoré nie sú kľúčovými, sa navyše ukladajú v zníženom rozlíšení. Pri popisovanom prístupe, je databáza tvorená blokmi kľúčových snímok. V článku [2] sa ako prístup pri následnom dopočítavaní využíva parameter energie, ktorú daný blok nesie, predstavujúcej hustotu detailu,

[1] Kým do roku 2018 bolo na Slovensku legálne uchovávať kamerové záznamy z bezpečnostných kamier po dobu maximálne 15 dní, v súčastnosti už je to len 72 hodín.



Obr. 1.2: Myšlienka ukladania každého k-teho záberu v pôvodnom rozlíšení z článku [2]

a bloky, ktoré prekročia určenú hranicu, sú nahradené prislúchajúcimi interpólovanými blokmi z databázy blokov kľúčových snímok, ktorá je o ne následne aktualizovaná.

1.2.1.2 Využitie rozsiahlej existujúcej databázy

Inou možnosťou je využívať rozsiahlu databázu "trénovacích"záberov, podobne ako pri trénovaní umelej inteligencie, a ku každému interpólovanému bloku v nižšom rozlíšení hľadať najpodobnejší blok vo vysokom rozlíšení v celej databáze, alebo jej časti určenej predošlou skúsenosťou, pre urýchlenie procesu. Výhoda tohto prístupu je, že bez ukladania akýchkoľvek snímok v pôvodnom rozlíšení umožňuje ešte značnejšie zníženie celkovej pamäte potrebnej pre uloženie rovnako dlhého záznamu. Naopak nevýhodou je, že napriek rýchlemu rozvoju v tejto oblasti, stále nemusí byť jednoduché dostať sa v rámci prijateľných nákladov k databáze dostatočnej robustnosti. Tento prístup je tiež náročnejší na následné prehľadávanie tejto databázy [11], čo plynie z veľmi veľkého množstva blokov vo vysokej kvalite, ktoré musí obsahovať, aby dokázala vyhovieť všetkým, alebo veľkej väčšine scenárov. Taktiež tento prístup nie je veľmi účinný pri identifikácii osôb, keďže je nepravdepodobné, že by sme človeka s danou tvárou našli v databáze.

1.2.2 Zvýšenie rozlíšenia prevzorkovaním

Každý záber videa zvyčajne obsahuje silno a slabo štruktúrované regióny. Silno štruktúrované regióny sa skladajú z takmer homogénnych oblastí obrazu, ako jednofarebné steny, obloha, či akýkoľvek iný celistvý zhluk pixelov, ktorý ľudské oko považuje za jeden ucelený objekt s jedným farebným odtieňom. Medzi týmito objektami sú v silno štruktúrovaných regiónoch veľmi jasne definované hranice (napríklad rám okna, na rozdiel od napríklad vĺn na vodnej hladine). Slabo štruktúrované regióny nie sú na

záberoch jednoznačne definované tvarmi a farbami. Dobrým príkladom sú napríklad ľudské vlasy, alebo aj črty ľudskej tváre. Existujú algoritmy, ktoré na základe týchto štruktúr a gradientov častí obrazu dopočítavajú pixely definujúce regióny vo vyššom rozlíšení. Zároveň sa osvedčilo spracúvať silno a slabo štruktúrované regióny separátne a v rôznych škálach pre rôzne priestorové štruktúry [11]. Algoritmy v tejto oblasti siahajú od jednoduchých, matematicky založených a pracujúcich v priestorovej doméne obrazu, využívajúcich napríklad interpolácie (ako [12]), po prístupy založené na príkladoch a učení na základe vzoriek [13]. Obe strany spektra nesú svoje problémy a hodia sa do odlišne náročných podmienok.

1.2.2.1 Matematicky motivované metódy

Najzákladnejšími matematickými metódami, určenými k prevzorkovaniu záberu, čiže aj zvýšenia rozlíšenia, bez akejkoľvek informácie na pozadí, kontextu, a bez súvislosti s ďalšími snímkami videa, sú interpolácie. Chýbajúce pixely dopočítavajú na základe už existujúcich v ich susednosti. Najpoužívanejšia a najefektnejšia interpolácia je bikubická, potom bilineárna, a technika najbližšieho suseda, ktorá je najrýchlejšou a najprimitívnejšou možnosťou. Napriek tomu, že interpolácie sú rýchlym riešením, neprinášajú dostatočne uspokojivé výsledky v doménach, kde je najvyšší dôraz kladený na jednoznačnosť zachytenej informácie [14]. Interpolácie, posudzujúce každú snímku videa individuálne, sú totiž principiálne obmedzené množstvom usmerňujúcej informácie v tej danej snímke a nedokážu zužitkovať kvantitatívnu výhodu videa oproti fotografiám [15]. Ďalším matematicky motivovaným prístupom je technika zvaná Maximum A Posteriori (MAP), založená na Bayesovskej teórii. Zvyšovanie rozlíšenia videa pomocou MAP techniky sa ukázalo úspešnejšie v zachovávaní ostrých rohov v obraze. Táto technika podlieha ďalšiemu výskumu, keďže kombinuje odhad obrazu vo vysokom rozlíšení a aproximujúcu sústavu rovníc spájajúcu ho so záberom z videa [14]. Podľa tohto článku sa ďalej oplatí predspracúvať zábery z videa pred tým, než sa pokúsime o zvýšenie ich rozlíšenia, a to najmä odstránením šumu a rozmazania. Rozmazanie je pre video inherentným problémom. Hoci ľudský mozog ho pri rýchlo sa striedajúcich záberoch potláča [15], pre algoritmy predstavuje značnú komplikáciu. Schopných algoritmov potláčajúcich rozmazanie bolo navrhnutých viacero: [3], [16], [17].

1.2.2.2 Zvyšovanie rozlíšenia s využitím neurónových sietí

Naproti matematickým metódam stoja prístupy založené na strojovom učení, ktoré donedávna v praxi čelili problémom s dostupnosťou dostatku trénovacích vzoriek, a stále do istej miery čelia, čo môže mať veľký dopad práve na unikátne detaily - obzvlášť tváre, čiastočne aj evidenčné čísla vozidiel. Ich vhodnosť pre praktické použitie je otázna, kvôli výkonu nutnému na rýchle fungovanie takýchto algoritmov pri pou-

Obr. 1.3: Výsledok (napravo) aplikácie prístupu [3] na rozmazaný záber (naľavo)

žití v reálnom čase. Po rozsiahlom prieskume sme došli k zisteniam, že v čase písania práce sú relevantné algoritmy dedikované zvyšovaniu rozlíšenia videa pomocou hlbokých neurónových sietí v stave tvorby zverejniteľnej implementácie [18] [19], alebo neposkytujú dostatočnú dokumentáciu k tomu, aby sme vrámci tejto práce vedeli tieto algoritmy otestovať na vlastných záznamoch z domény bezpečnostných kamier [20] [21] [22], alebo zdrojový kód vôbec nezverejňujú [23]. Veľmi sľubný, rozsiahly a pomerne nový toolkit, obsahujúci množstvo algoritmov zameraných na prepojenie počítačového videnia a hlbokých neurónových sietí, ktoré obsadili popredné miesta na udalosti NTIRE (https://cvlai.net/ntire/2024/), vrátane zvyšovania rozlíšenia videa s využitím viacerých dedikovaných modelov a algoritmov, je MMagic [24]. Žiaľ však neposkytuje dostatočne obsiahlu dokumentáciu, a návody k použitiu metód relevantných pre našu prácu sú momentálne k dispozícii len v čínskom jazyku, z ktorého automatický, aplikáciami tretích strán sprostredkovaný preklad do angličtiny je pri odbornej terminológii veľmi nespoľahlivý.

1.2.3 Zlepšenie separovania odtieňov

Prístupy zvyšujúce rozlíšenie napomáhajú zvýrazneniu detailov na zábere, samostatný význam však nesú len za dobrých svetelných podmienok. Za zníženého osvetlenia je prvoradé objekty a tvary na zábere viditeľne separovať odtieňmi. Je ľahké presvedčiť sa pomocou očných klamov, že zrakové centrum ľudského mozgu vníma farebné odtiene relatívne, porovnáva ich medzi sebou navzájom. Robí tak, aby za ľubovoľných svetelných podmienok zachovalo schopnosť odhadnúť skutočnú farbu objektov a klasifikovať ich na jej základe. V digitálnej podobe však pixelom priraďujeme absolútne hodnoty z určitého diskrétneho rozsahu, ktorý je ešte k tomu značne menší, než rozsah odtieňov a jasností, s akými vie pracovať ľudský zrakový systém [8]. V rámci tohto digitálneho

rozsahu by sme sa preto chceli čo najviac priblížiť takému podaniu vizuálnej informácie, na základe ktorého dokáže mozog spoľahlivo separovať a klasifikovať objekty. Zlepšenie záberu vieme robiť na základe jeho samotného, alebo na základe ďalších záberov. Medzi najtradičnejšie spôsoby samozlepšenia snímku patrí zlepšenie kontrastu. Metódy zameriavajúce sa na zlepšenie kontrastu sa delia na priame – tie definujú mieru kontrastu a snažia sa ju zlepšiť, a nepriame – tie zlepšujú kontrast využitím nevyužitých častí dynamického rozsahu bez toho, aby definovali nejakú špecifickú mieru kontrastu [7]. Algoritmy zlepšujúce kontrast využívajú dva základné prístupy. Prvým je ekvalizácia histogramu, kedy sa obvykle snažíme čo najrovnomernejšie rozprestrieť početnosti pixelov jednotlivých intenzít po celom spektre. Nemusíme sa však snažiť len o rovnomerné rozdelenie, využíva sa aj logaritmický prístup, alebo power-law pravidlo [25]. K zachovaniu úrovní vstupných jasov bolo navrhnutých viacero metód založených na ekvalizácii histogramu: Stred zachovávajúca bi-histogramová ekvalizácia (BBHE - mean preserving bi-histogram equalization), rovno-obsahová dualistická podobrazová ekvalizácia (DSIHE - equal area dualistic sub-image histogram equalization), bi-histogramová ekvalizácia minimalizujúca strednú jasovú odchýlku (MMBEBHE - minimum mean brightness error bi-histogram equalization), rekurzívna stredno-rozptylová histogramová ekvalizácia (RMSHE - recursive mean-spread histogram equalization) a multi-histogramová ekvalizácia [7]. Druhým zo základných prístupov, ktoré využívajú algoritmy zlepšujúce kontrast záberov, je mapovanie odtieňov. Jeho cieľ je v podstate rovnaký, ako pri ekvalizácii histogramu – komprimovať alebo premapovať dynamický rozsah HDR obrazov do rozsahu vhodného pre zobrazenie na zariadeniach so štandardným dynamickým rozsahom (SDR) [8], zachovajúc pri tom čo najviac detailov a vizuálnej informácie a pokiaľ možno pôsobiac prirodzene – hoci prirodzený vzhľad obrazu nemusí byť pri bezpečnostných kamerách prioritou (má však význam napríklad aj pre redukciu zrakovej a kognitívnej únavy v prípade, že za monitormi systému dlhodobo sedí poverený pracovník). Medzi algoritmy premapovania odtieňov patria napríklad nasledovné techniky:

- Reinhardove mapovanie odtieňov sa snaží napodobniť spôsob, akým sa ľudský zrak adaptuje na rozličné svetelné podmienky. Komprimuje dynamický rozsah obrazu tak, aby zachovalo lokálny kontrast a saturáciu. Pôvodne išlo o asistovaný proces - vyžadoval od zhotoviteľa snímky vyznačenie najsvetlejšieho, najtmavšieho a stredne tmavého (referenčného) miesta v obraze k vypočítaniu dynamického rozsahu. Neskôr Reinhard a kolektív automatizoval tento proces využívajúc kruhový Gaussov operátor v rôznych škálach aplikovaný na celý záber, opravujúc presvetlené a tmavé regióny. [26]

- Fattalove mapovanie odtieňov taktiež priorizuje lokálny kontrast a snaží sa ho zlepšiť, zatiaľ čo zábery si zachovávajú prirodzene pôsobiaci vzhľad. Komprimuje gradient jasového komponentu obrazu a riešením jeho Poissonovej rovnice

konštruuje prislúchajúci obraz v zobraziteľnom dynamickom rozsahu. [27]

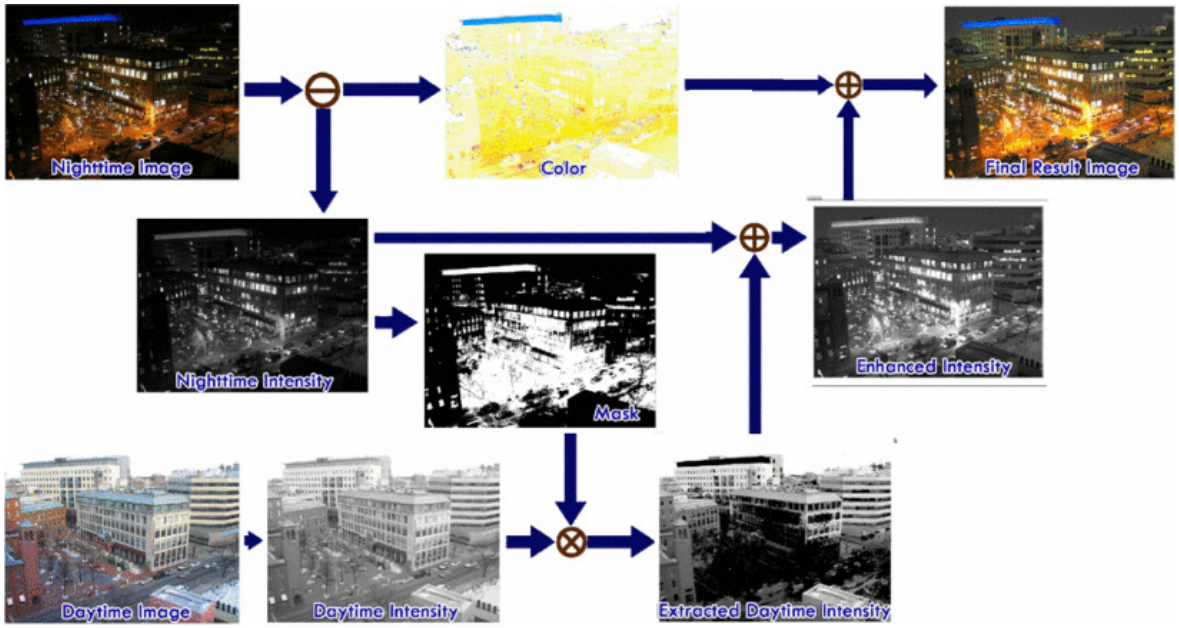
- Dragove mapovanie odtieňov sa opäť zameriava na lokálny kontrast, a taktiež na potlačenie halo efektov, ktoré sú bežným nežiadúcim sprievodným javom premapovávania odtieňov. Pracuje s adaptívnym upravovaním logaritmických základov, na báze ktorých sa komprimujú hodnoty jasnosti v obraze. Na zlepšenie tmavých oblastí používa zvyšovanie kontrastu. [28]
- Mantiukove mapovanie odtieňov modeluje generický operátor, ktorý aproximuje lokálne a globálne operátory využité v iných algoritmoch výpočtovo nenáročnými operáciami. Jeho výsledky sú často vizuálne neodlíšiteľné od algoritmov na mapovanie odtieňov vyžadujúcich značné množstvá strojového času. [29]

Ukázalo sa, že exponenciálne operátory mapovania odtieňov prinášajú vyššie subjektívne vnímané zlepšenie dynamického rozsahu než logartimické [8]. Mapovanie odtieňov sa obvykle vykonáva len v kanáli iluminácie a na logaritmickej škále, pričom mapa žiarenia vo floatovej reprezentácii sa konvertuje do 8-bitovej reprezentácie použiteľnej k renderovaniu [7]. Vyšší dynamický rozsah možno taktiež dosiahnuť kombináciou viacerých obrázkov z rovnakej scény s rôznymi časmi expozície. Relatívne intenzity presvetlených častí sú kvalitne zachytené pri krátkej uzávere objektívu, zatiaľ čo pri dlhej stihne na šošovku dopadnúť dostatok svetla na to, aby vynikli aj rozdiely medzi intenzitami v tmavej časti záberu [8]. Výsledkom je mapa floatov reprezentujúcich hodnoty žiarenia s hodnotami proporčnými tomu v skutočnej scéne. Tento spôsob je však v kontexte bezpečnostných kamier prakticky takmer nepoužiteľný, keďže je nutná podpora zo strany samotnej kamery, ktorá nie je u väčšiny produktov samozrejmosťou, a to z logického dôvodu, že luxus vyhotovovania jedného záberu niekoľkokrát sa veľmi ťažko zmysluplne realizuje v rámci videa. V tom obvykle priorizujeme počet snímok za sekundu, nehovoriac o výpočtovej náročnosti okamžitého kombinovania a ladenia týchto záberov. To je výzvou napríklad aj pre najmodernejšie vlajkové lode medzi súčasnými smartfónmi, využívajúcimi vysokovýkonné čipy, akými bezpečnostné kamery v praxi nedisponujú, a z ekonomických dôvodov pravdepodobne ešte dlho disponovať nebudú.

1.2.4 Kontextová fúzia

Algoritmy zlepšujúce všeobecné videá zvyčajne venujú zvýšenú pozornosť odstraňovaniu zahmlenia, šumu, oparu a atmosférických javov. Neprodukujú však uspokojivé výsledky pri nočných záznamoch z bezpečnostných kamier, pretože sú závislé od jasu pixelov [4]. Výhodou domény bezpečnostných kamier, ktorú uchopili niektoré algoritmy, je, že kamery sú obvykle umiestnené staticky pred tú istú scénu po celý čas – teda cez deň aj v noci. Tým pádom si dokážeme ukladať referenčné zábery, na ktorých

je značná časť scény zhodná so zábermi vyhotovenými v ľubovoľnej fáze dňa, za ideálnych svetelných podmienok, a použiť tieto na dopĺňanie informácii vysokej kvality za zníženej viditeľnosti. Tomuto procesu sa najčastejšie hovorí kontextová fúzia a algoritmus kombinuje snímky z rôznych častí dňa automaticky. Jednu z možných schém f[zie] vidíme na obrázku 1.4



Obr. 1.4: Schéma kontextovej fúzie dát z denného a nočného záberu scény navrhnutá v článku [4]

Algoritmy na fúziu možno rozdeliť do kategórii na spodnú, strednú a vysokú úroveň. Tieto sa tiež označujú ako úroveň pixelov, úroveň čŕt a symbolická úroveň [7]. Za základný môžeme považovať prístup, kedy detekujeme hrany v obraze, a na základe nich ho rozdelíme do segmentov, pre ktoré sa snažíme určiť správny odtieň spárovaním s prislúchajúcimi segmentmi v záberoch za lepších svetelných podmienok. Veľkým prínosom pre detekciu hrán postáv, ktoré patria obvykle medzi najzaujímavejšie informácie vo videu, vie byť infračervené spektrum. Mnohé bezpečnostné kamery disponujú aj týmito infračervenými senzormi. Očakávateľne však nastáva značná komplikácia pri rekonštrukcii odtieňov objektov, ktoré sa za svetla pred kamerou nevyskytovali, a odhad ich farby, ktorá je relatívna voči objektom, ktoré sa nachádzali aj na svetlom aj na tmavom zábere, je výzvou. Ďalším stávajúcim problémom je ako sa vysporiadať s neuniformným osvetlením. Mnohé z algoritmov sú založené na uniformnej iluminácii, a nevedú si dobre ak máme viacero rôznorodých zdrojov svetla. Tými sú zvyčajne jeho umelé zdroje ako lampy, reflektory, ich odrazy a podobne. V takýchto podmienkach je esenciálna správna klasifikácia osvetlených a tmavých regiónov [4]. Najsľubnejšie sa v tomto ohľade javí nasadenie neurónových sietí, napríklad v podobe samo-organizujúcich sa máp.

1.3 Súhrn východísk

Vzhľadom na rozsiahlosť oblastí, v ktorých sa žiada zlepšovať kvalitu digitálneho obrazu, a ich individuálnu komplexnosť, nebol doposiaľ vyvinutý algoritmus, ktorý by kombinovane cielil k zlepšeniu každej, alebo väčšiny zo zmysluplných metrík kvality. Zámerom tejto práce bude preskúmať kombinácie algoritmov pre subdoménu bezpečnostných kamier, avšak takých, ktoré sú v stave nasaditeľnosti do prevádzky, alebo ich sprevádzkovanie je v súlade s rozsahom tejto práce. Mnohé z algoritmov prezentovaných v tejto kapitole žiaľ neponúkajú žiaden zdrojový kód ani prevzateľný softvér, a nebudeme ich teda môcť otestovať. Viac o výbere testovaných algoritmov sa čitateľ dozvie v podkapitole 2.4.

Kapitola 2

Ciele a metodika práce

V tejto kapitole sa pozrieme na to, ako budeme postupovať pri aplikovaní akých prístupov a algoritmov na aké testovacie kamerové záznamy.

2.1 Čas nasadenia algoritmu

Vzhľadom na nezanedbateľnú časovú náročnosť niektorých operácii nad snímkami hrá rozdielovú rolu kedy v procese zaznamenania, spracovania, prípadného uloženia a zobrazenia záberu, sa aplikuje algoritmus.

- Pri systémoch, ktoré vyhotovovaný záznam ukladajú bez toho, aby bol v reálnom čase sledovaný ľudským strážnikom, alebo nejakým algoritmom vyhodnocujúcim kontext a potenciálnu nežiadúcu aktivitu v priestore, a teda tento záznam bude vždy prehliadaný až po tom, ako došlo k aktivite, ktorú máme záujem na ňom preskúmať, nie sme obmedzení časom na spracovanie videa. Aspoň nie do značnej miery. Môžeme si teda dovoliť napríklad 10 sekundový úsek videa, povedzme že tvorený 10 sekúnd $\times$ 30 fps = 300 snímkami, spracúvať pokojne niekoľko minút. Nie je problematické spracúvať jednu snímku aj rádovo v desatinách sekundy. Zároveň nečelíme žiadnym kompromisom v prípade nasadenia takého algoritmu, ktorý na zlepšenie záberu v čase t používa informácie zo záberu v čase $t+1$ alebo neskoršieho.

- Ak naopak zábery zobrazujeme v reálnom čase, je žiadúce, aby tieto zábery už boli zlepšené, a teda zvýrazňovali informácie dôležité pre strážnika (či už v roli fyzickej osoby alebo softvéru). V takom prípade je čas na zlepšenie snímky limitovaný počtom záberov snímaných za sekundu. V prípade algoritmov využívajúcich neskoršie zábery na zlepšenie predošlých, sa musíme zmieriť s oneskorením. To je však pri vhodne zvolenom algoritme konštantné a nemusí presahovať rádovo desiatky sekúnd, ktoré málokedy predstavujú rozdiel medzi dovolením a predídením

nežiaducej udalosti (ale existujú aj také prípady nasadenia, kedy o účinku aktivácie bezpečnostných opatrení - obvykle na diaľku - rozhodujú tieto potenciálne desiatky sekúnd).

V tejto práci sa pri testovaní a posudzovaní algoritmov najprv zameriame na prvý prípad, a teda nebudeme brať do úvahy časovú limitáciu na zlepšenie záberov. Následne pri prístupoch, ktoré sa ukážu ako efektívne, otestujeme alebo prepočítame na základe času, ktorý sme odmerali, že potrebujú, či by bolo možné ich (poprípade v nejakej odľahčenej, kompromisnej forme) aplikovať aj na druhý prípad, a zlepšovať nimi video v reálnom čase.

2.2 Testovacie dáta

Hoci mnohé z algoritmov budú nezávislé od kontextu videa a záberov iných než práve spracúvaného, nebudeme ich testovať na individuálne stojacich snímkoch, ale rovno na videách. Tieto videá zaobstaráme z niekoľkých zdrojov, aby sme pokryli čo najviac realistických situácii:

- Za pomoci reálnych bezpečnostných kamier zapožičaných školou sme nahrali niekoľko vlastných záznamov v relevantých prostrediach a za rôznych svetelných podmienok, z rôznych vzdialeností. Aby nedošlo k porušeniu niečej ochrany súkromia, postavou pohybujúcou sa v popredí záberov bol autor tejto práce. Prvá z kamier je Annke T200. Táto kamera popri RGB zázname vyhotovuje aj termálny záznam, ten však nie je v rovnakom rozlíšení, ani nezaberá rovnaký výsek obrazu ako ten z RGB senzoru, preto bude nutné k jeho využitiu na zlepšenie prislúchajúceho RGB videa nájsť zodpovedajúcu homografiu (viac v časti 2.4.7). Touto kamerou sme nahrali dve dvojice prislúchajúcich RGB a termovíznych záberov, jednu za dňa, a druhú večer po západe slnka, na tom istom mieste. Na týchto videách v rozlíšení 1920×1080 pixelov na RGB zázname (interne nazvaných Annke_RGB_day a Annke_RGB_dark) a v rozlíšení 320×240 pixelov na termo zázname (interne nazvanom Annke_thermo_day a Annke_thermo_dark), sa prechádza autor vo vzdialenosti asi 5 až 25 metrov od objektívu kamery po parkovisku a vo večernom videu aj na priľahlom trávniku, viď ukážky 2.1a, 2.1b, 2.2a, 2.2b.

  Druhá z kamier je Hikvision DS-2CD2025FWD s infračerveným prisvecovaním. Túto kameru bolo obtiažnejšie spojazdniť, no nakoniec sme prišli na to, že elektrický zdroj (v balení originálny dodaný nebol) mal nepatrne väčší priemer konektoru než kamera, a teda kamera ním nebola napájaná. Problém bol provizórne vyriešený vsunutím drôtika a kameru sme uchytili k doske, ktorú sme pripevnili k

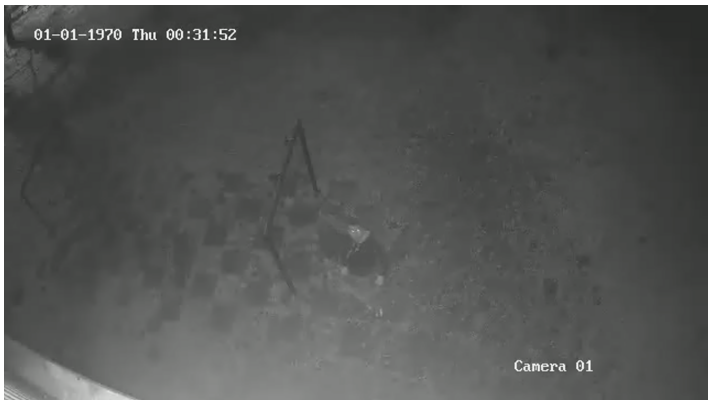
(a) Annke_RGB_day (b) Annke_thermo_day

Obr. 2.1: Ukážky testovacieho záznamu z kamery Annke T200 za dňa

(a) Annke_RGB_dark (b) Annke_thermo_dark

Obr. 2.2: Ukážky testovacieho záznamu z kamery Annke T200 večer

zábradliu na balkóne u autora práce doma, na druhom poschodí. Ten sa následne išiel poprechádzať v zornom poli kamery, a to najprv vo vzdialenosti asi 8 metrov vzdušnou čiarou pri pohľade kamery takmer kolmo dole vo večernej tme na videu Hikvision_dark_close v rozlíšení 640 × 360 pixelov 2.3, a potom vo vzdialenosti zhruba 25 až 40 metrov od objektívu kamery v priestore prisvietenom pouličnou lampou na videu Hikvision_dark_distant 2.4, v rozlíšení orezanom z pôvodných 640 × 360 pixelov na 480 × 360 pixelov.



Obr. 2.3: Ukážka videa Hikvision_dark_close

Obr. 2.4: Ukážka videa Hikvision_dark_distant

- K prekvapivému zisteniu sa na základe nášho prieskumu na internete nenachádza verejne dostupný dataset záberov z bezpečnostných kamier. Takýto mal byť k dispozícii zdarma na stránke https://viratdata.org/, avšak prístup k nemu už zjavne neexistuje. Na stránke sú však dve ukážkové videá, ktoré mal dataset obsahovať. Prvé video, pre účely práce pomenované student_street, je v rozlíšení 1280 × 720 pixelov a zachytáva zhruba z výšky prvého nadzemného podlažia verejné vonkajšie priestory pred budovami patriacimi univerzite v Spojených Štátoch Amerických, v ktorých sa pohybujú ľudia za polooblačného počasia počas dňa, vo vzdialenosti asi 10 až 50 metrov od objektívu kamery (viď ukážka 2.5).

Obr. 2.5: Ukážka videa student_street

Druhé video, pre účely práce pomenované "parking", je v rozlíšení 1920 × 1080 pixelov a zaznamenáva parkovisko v zadných priestoroch rozmernej budovy so

Obr. 2.6: Ukážka videa parking

zaparkovanými vozidlami, v daždi počas dňa, z výšky asi tretieho až štvrtého nadzemného podlažia (viď ukážka 2.6). Na 53 sekúnd trvajúcom zázname prejde 5 ľudí s dáždnikmi vo vzdialenosti odhadom 20 až 60 metrov od objektívu kamery. Kamera vykazuje mierne pohyby následkom prítomného vetra.

- Existuje množstvo verejných kamier dohľadového alebo podobného typu, z ktorých záznamy sú naživo zdieľané na internete, zadarmo a dostupne pre širokú verejnosť. IP adresy týchto kamier možno získať zo zdrojových kódov webstránok, na ktorých sa nachádzajú, a napojiť sa na ne cez HTTP, alebo iný protokol pomocou nami napísaného programu. Z týchto záznamov možno uložiť ich časť, alebo na ne aplikovať algoritmy naživo, počas streamovania. Kvalita týchto videí je rôzna, väčšinou však trpí nízkym počtom snímok za sekundu, nízkym rozlíšením, a vysokou mierou šumu, najmä v noci. Tieto zábery sú tak dobrou výzvou pre algoritmy, napriek tomu, že sú spravidla horšej kvality, než tie vyhotovené väčšinou moderných bezpečnostných kamier. Ich výhoda je, že sú zbierané z kamier umiestnených naprieč celým svetom, a teda v rôznych časových pásmach a za rôznych poveternostných podmienok. Obvykle ale sledujú prostredie z väčšej vzdialenosti, než je bežné pre bezpečnostné kamery, a niekedy ide o prostredia, v ktorých by nasadenie bezpečnostných kamier ani nebolo opodstatnené.

  Ako súčasť implementácie sme si vytvorili pomocný modul live_video_acquirement.py s triedou LiveVideo, obsluhujúcou živé pripojenie v reálnom čase na vybrané kamery z rôznych prostredí. Táto trieda jednak poskytuje funkcionalitu na nahratie úsekov videa z týchto verejných kamier, jednak spravuje ich zoznam, a zároveň bude slúžiť hlavnej triede, v ktorej sa aplikujú algoritmy na video, aby sme ich vedeli nasadiť priamo na živý prenos.

Obr. 2.7: Ukážka videa american_crossroad_traffic

Pomocou triedy LiveVideo sme zhotovili 4 záznamy, na ktorých budeme testovať rôzne metódy zlepšenia. Na zázname american_crossroad_traffic (viď ukážku 2.7) je kolóna áut stojacich privrátene ku kamere, kým v opačnom smere prúdi od kamery premávka. Toto video má rozlíšenie 800 × 450 pixelov a v najbližšom bode sú autá asi 10 metrov vzdušnou vzdialenosťou od kamery. Vzhľadom na nízku kvalitu záznamu, ktorý ani nie je originálne vyhotovovaný za účelom monitorovania bezpečnostným systémom a identifikácie osôb a vozidiel, nepredpokladáme, že sa nám podarí zaznamenať evidenčné čísla vozidiel. Napriek tomu sa na to môžeme sústrediť, pretože už malé obohatenie informácie môže vyvrátiť výskyt, alebo zvýšiť pravdepodobnosť výskytu vozidla s danou štátnou poznávacou značkou v danom čase na danom mieste. Nižšie rozlíšenie než na bežných bezpečnostných kamerách tiež nie je irelevantné, keďže stačí, aby bola kamera s vyšším rozlíšením umiestnená ďalej od pozorovaných vozidiel a efekt (rozlíšenie vizuálnej informácie) by bol rovnaký.

Záznam dutch_street_sunset (viď ukážka 2.8) bude výzvou v inom ohľade než rozlíšení. Ide o záber ulice v obytnej zóne, po ktorej prejde auto, približujúce sa a vzďalujúce sa od kamery. Tento záznam v rozlíšení 1280 × 720 pixelov je však vyhotovený počas západu slnka pri jasnej oblohe. Slnko je zachytené v ľavej časti obrazovky a ukazuje obrovskú limitáciu dynamického rozsahu. Na zábere je celá ulica, privrátená tienistou stranou, veľmi tmavá, zatiaľ čo obloha je presvetlená západom slnka.

Na zázname german_marketplace_eve (viď ukážka 2.9) je z výšky zhruba piateho nadzemného podlažia zachytené námestie, na ktorom sa pohybujú ľudia, z toho jeden človek prejde skrz na bicykli. Ide o večernú scénu, kde časti námestia sú zaliate svetlom pouličných lámp, zatiaľ čo iné časti sú podstatne tmavšie.

Obr. 2.8: Ukážka videa dutch_street_sunset

Vzhľadom na rozlíšenie 800 × 600 pixelov a vzdialenosť od pohybujúcich sa osôb, nie je šanca rozpoznávať tváre, bezohľadu na sofistikovanosť algoritmu. Nemiesto toho sa zameriame najmä na rozoznateľnosť detailov vo videu.

Obr. 2.9: Ukážka videa german_marketplace_eve

Záznam strba_snow_fog (viď ukážka 2.10) v rozlíšení 1280 × 720 pixelov neobsahuje žiadne pohybujúce sa objekty, avšak samotná kamera sa pomaly otáča pozdĺž horizontálneho oblúku. Toto video je zaujímavé z toho hľadiska, že je na ňom športový areál zachytený v hmle a snehu. Toto je poveternostný scenár, ktorý je bežný v zimných mesiacoch, no nie v mesiacoch písania práce. Preto

Obr. 2.10: Ukážka videa strba_snow_fog

sme si vybrali tento areál na otestovanie správania a prípadných nedostatkov algoritmov v takýchto špecifických, no nie neobvyklých podmienkach.

Okrem nahraných videí, budeme vedieť testovať aj prístup k týmto a zopár ďalším streamom v reálnom čase, a zlepšovanie kvality ich záberov bez spôsobenia omeškania a signifikantného výpadku nespracovaných záberov. Medzi ďalšími kamerami, na ktoré sa nám podarilo pomocou OpenCV pripojiť, je napríklad kamera na pláži v Brazílii, situovaná pred plastovými stolami a stoličkami v popredí paliem, alebo kamera na malom lokálnom letisku, v Spojených Štátoch Amerických, obsluhujúcom veľké množstvo malých motorových lietadiel.

Problémom, ktorému ale čelíme, a ktorý sa nám nepodarilo vyriešiť, je nízka frekvencia snímok za sekundu získaných po pripojení sa na vzdialené kamery cez Python. Na webstránkach, kde tieto scenérie možno sledovať naživo, je počet snímok za sekundu vyšší, než k akému vieme pristupovať v rámci nášho programu. Hlásená hodnota počtu snímkov za sekundu prenášaného videa je však pôvodná, nezodpovedajúca snímkam prijatým na našom konci, a teda týchto menej snímok naskladaných do záznamu s rovnakou frekvenciou zapríčiňuje zrýchlený chod nahraného videa. Vzhľadom na podstatu testovaných algoritmov však nejde o kritický nedostatok.

2.3 Návrh testovacieho programu

Implementačným jadrom tejto práce bude program, pomocou ktorého budeme môcť na testovacie videá aplikovať rôzne zlepšovacie algoritmy, kombinovať ich, a porovnať ich výsledky v subjektívnych (napr. vizuálny posudok) a objektívnych (napr. čas aplikácie

prepočítaný na jednu snímku v závislosti od jej rozlíšenia, alebo jeden pixel, prípadne iných metrík) mierach.

2.3.1 Voľba programovacieho jazyka a knižníc

Tento program bude napísaný v programovacom jazyku Python, verzia 3.12. Rozhodli sme sa tak preto, lebo pre Python existuje solídna paleta open-source knižníc, ktoré ponúkajú rozsiahle schopnosti v oblasti práce s obrazom a digitálnou informáciou, a majú efektívne implementovaných množstvo funkcii a algoritmov relevantných pre našu prácu. Medzi použité knižnice. bude okrem vstavaných modulov a modulov prevzatých z githubových stránok algoritmov, patriť:

- OpenCV, verzia 4.9, je Intelom vyvíjaná open-source knižnica obsahujúca viac než 2500 algoritmov v oblasti počítačového videnia,
- NumPy, verzia 1.26, poskytuje infraštruktúru pre prácu s vektormi a maticami (ktorými je digitálny obraz reprezentovaný), a funkciami nad nimi definovanými,
- PIL (Python Imaging Library), verzia 10.2.0, je open-source knižnica slúžiaca na manipuláciu s digitálnym obrazom,
- PyTorch, verzia 2.2.1+cpu, je knižnica strojového učenia založená na knižnici Torch.

Hoci jazyk C++, s ktorým taktiež funguje knižnica OpenCV, je všeobecne rýchlejší než Python, keďže je kompilovaný, a teda transformovaný priamo do strojového kódu, OpenCV v Pythone je len wrapperom nad originálnym kódom v C++. Kombinuje tak rýchlosť C++ a jednoduchosť Pythonu. Rozdiel v rýchlosti behu OpenCV algoritmov medzi C++ a Pythonom je obvykle menej ako 1%, a v najhoršom prípade, pre základné funkcie je uvádzaný ako menej než 4%. Táto úvaha je dôležitá kvôli tomu, že práca s obrazom je všeobecne veľmi výpočtovo náročná, o to viac, keď sa pozrieme napríklad na hlboké neurónové siete nad obrazovými dátami a budeme ich chcieť využiť na úpravu záberov vo videu.

2.3.2 Porovnávanie výsledkov

- Pri algoritmoch, neporovnávajúcich rozlíšenie pôvodného a upraveného záznamu, si tieto videá zobrazíme paralelne v jednom rozdelenom okne, bikubickou interpoláciou naškálované tak, aby toto okno zaberalo celú obrazovku monitora.

- Pri algoritmoch zameraných (aj) na rozlíšenie, sa zameriame na takú časť obrazu, ktorá po preškálovaní zaberie zhruba polovicu monitora, aby bolo možné najjednoznačnejšie porovnať výsledok s preškálovaním rovnakého výseku na rovnakú veľkosť inou formou.

Na zobrazovanie videí vytvoríme prehrávač, ktorý dokáže prehrať jedno až štyri videá na jednom monitore naraz, a to tak, aby maximalizoval priestor, ktorý tieto videá vypĺňajú, v prospech jednoduchšieho vizuálneho porovnania kvality. Ak nebudeme chcieť videá o rovnakej veľkosti škálovať, k dispozícii budú aj sesterské metódy, ktoré dodržia pri prehrávaní pôvodné rozlíšenie. Tento prehrávač taktiež umožní tieto koláže paralelne bežiacich videí ukladať v nejakom zo štandardných formátov ako mp4 alebo avi.

Zároveň do programu zahrnieme checkpointy, v ktorých bude zaznamenaný čas od spustenia programu, aby sme vedeli vyhodnotiť a porovnať časovú náročnosť jednotlivých procedúr. Program bude tiež obsahovať zabudovanú infraštruktúru umožňujúcu dostatočne rýchle procedúry aplikovať v reálnom čase na živé záznamy z online kamier.

2.3.3 Aplikácia algoritmov

Každý z testovaných algoritmov bude implementovaný v samostatnej metóde, ktorá bude umožňovať nastavenie jeho parametrov (ak nejakými disponuje) pri jej volaní. Vstupom do týchto metód aj výstupom z nich bude video, takže bude možné takto aplikáciu algoritmov reťaziť. Okrem toho pre algoritmy aplikované na zábery samostatne, väčšinou tie, špecializujúce sa všeobecne na zlepšovanie nejakej z metrík kvality digitálneho obrazu, ktoré sme si zaviedli v prvej kapitole, vytvoríme metódy, ktoré budú tieto procedúry aplikovať len na jeden vstupný záber a vracať tento záber upravený. To sa nám bude hodiť, keď budeme chcieť aplikovať kombináciu zlepšovacích algoritmov záber po zábere, napríklad pri spracúvaní videa v reálnom čase. Nastavenia parametrov týchto algoritmov budú mať globálnu pôsobnosť a meniteľnosť.

2.4 Testované algoritmy

Otestujeme relevatné algoritmy, ktoré sú obsiahnuté v open-source knižniciach, alebo ktorých kód je zverejnený na internete. Pokiaľ je nasadenie algoritmu komplikované, budeme pracovať len s takými kódmi a knižnicami, ktoré majú prehľadnú dokumentáciu v angličtine. Ťažiskom tejto práce nie je vývoj a zavádzanie procedúr na základe teoretických základov, je ním použitie a porovnanie širokého množstva metód v doméne bezpečnostných kamier. Taktiež tým pádom opomenieme experimentálne návrhy algoritmov testovaných len na akademickej pôde, bez implementácie použiteľnej v reálnej prevádzke na vlastné účely, napriek tomu, že sme takéto algoritmy spomenuli v

prvej kapitole, keďže predstavujú zaujímavé nové možnosti, ktoré pri rýchlosti vývoja technológii digitálneho obrazu môžu byť už v dohľadnej dobe jednoho až pár rokov zavádzané do prevádzky.

Nešťastným, no pochopiteľným faktom, ktorý znižuje strop možností tejto práce, je, že v podstate veštky najschopnejšie moderné implementácie softvérového skvalitňovania digitálneho obrazu sú v rukách globálnych technologických spoločností, ktoré si starostlivo strážia svoje know-how, do ktorého získania na výskumnej úrovni, a následne jeho vyladenia pre praktické nasadenie, investovali milióny dolárov. Open-source algorimty, prezentované malými univerzitnými a nadšeneckými tímami teda len ťažko dokážu byť kompetitívne proti technológiam zabudovaným do zariadení od veľkých svetových výrobcov elektroniky, medzi ktorými panuje stav ostrej konkurencie a rivality, a na ktorých sa podieľalo mnohonásobne viac odborníkov s mnohonásobne väčšími dostupnými prostriedkami. Táto rivalita je v posledných rokoch najväčšia na trhu smartfónov, kde spoločnosti ako Apple, Samsung, Huawei, Xiaomi a ďalšie dokázali vyvinúť pre svoje kamerové sety v telefónoch softvér, ktorý silou dodatočného spracovania často dorovnáva dedikované digitálne fotoaparáty a zrkadlovky napriek tomu, ako veľmi sú znevýhodnené fyzikálnymi limitmi niekoľkonásobne menšej fotografickej aparatúry. Toto platí aj v porovnaní s dedikovanými kamerami, o to viac, keď sa zaoberáme bezpečnostnými kamerami, ktoré sú jednak cenovo značne limitované na úrovni individuálneho zariadenia, a zároveň na úrovni spoločností, keďže toto odvetvie nie je dostatočne ziskové na to, aby si výrobcovia bezpečnostných kamier mohli dovoliť vývoj softvérov dodatočného spracovania, aký si môžu dovoliť poprední výrobcovia smartfónov. To nás však neodrádza, prejdime si teda algoritmy, ktoré dokážeme implementovať v tejto práci:

2.4.1 Zaostrenie obrazu

Pri zaostrovaní obrazu využijeme dve metódy:

- Aplikáciu Laplaceovského hranového filtra, kde otestujeme rôzne matice, ktorých prikladaním ku každému pixelu a jeho okoliu zvýrazníme hrany v zábere.

- Zaostrenie odčítaním rozmazaného obrazu, kedy si najprv vytvoríme kópiu záberu, tú rozostríme Gaussovským rozmazaním a 0.X násobok tohto rozostreného obrazu odčítame od 1.X násobku pôvodného obrazu, aby sme zachovali rozsah intenzít, pričom vyskúšame rôzne hodnoty X, ako aj rôzne hodnoty okolia a strednej odchýlky pri rozmazávaní.

Obe metódy sú všeobecne dlho známe a používajú sa pri zaostrovaní digitálnych obrazov, bez ohľadu na doménu. Nasádzať ich teda budeme snímku po snímke a uvidíme,

Obr. 2.11: Ukážka výsledku odstraňovania šumu nelokálným priemerovaním z článku [5]. Naľavo je pôvodný obraz, uprostred obraz zašumený Gaussovským šumom a napravo je obraz po aplikácii algoritmu. Všimnime si, že táto fotografia je veľmi vhodná na jeho nasedenie, pretože obsahuje veľa opakujúcich sa častí, priemerovaním medzi ktorými možno spoľahlivo odstrániť šum.

aký efekt budú mať na videá z bezpečnostných kamier cez deň, v noci, a v rôznych iných podkladových atribútoch kvality.

2.4.2 Odstránenie šumu

Nežiadúcim efektom pri videu z bezpečnostnej kamery by bolo, keby sme pri snahe odstrániť šum rozmazali obraz vyhladzovaním a nevyberane ho tak zbavili informácie. V našej oblasti je totiž zachytená informácia prioritou, a uhladený vzhľad videa je až sekundárny - ani nie tak pre ľudských pozorovateľov, ako pre ďalšie algoritmy, ktorých výsledok nasadenia by mohol šum znehodnotiť, ako je napríklad veľmi reálne pri metódach využívajúcich hlboké neurňové siete, ktoré v záujme lepšieho výsledku preferujú čo najmenej rušivej šumovej informácie nesúvisiacej s kontextom. Preto použijeme metódu odstraňovania šumu nelokálným priemerovaním [5]. Jej algoritmus je implementovaný v OpenCV a pozerá sa na obraz po malých výrezoch o veľkosti rádovo v pixeloch, pre ktoré hľadá výrezy rovnaké alebo takmer rovnaké inde v obraze (pre lepšiu predstavu viď obr. 2.11), a následne priemeruje hodnoty ich pixelov, aby tak z nich odstránil jedinečný šum. Táto metóda je značne pomalšia než jednoduché metódy vyhladzujúce obraz ako napríklad Gaussovské rozmazanie, no kompromituje informáciu obsiahnutú v zábere do podstatne nižšej miery. Jej sila zároveň rastie, ak využijeme k vyhľadávaniu podobných častí obrazu aj ďalšie snímky, ktoré máme v prípade videa k dispozícii. To je obzvlášť prínosné v prípade bezpečnostných kamier, kde nás často zaujíma práve informácia nevyskytujúca sa nikde inde na zábere - napríklad tvár konkrétneho človeka - avšak obsiahnutá na ďalších záberoch, a to s vysokou pravdepodobnosťou nezašumená, alebo aspoň zašumená v iných bodoch inou intenzitou.

2.4.3 Zlepšenie kontrastu

O zlepšenie kontrastu v testovacích videách (tam, kde sa to bude javiť ako prínosné), sa pokúsime pomocou ekvalizácií histogramu. Budeme ich aplikovať na jednotlivé snímky individuálne. V špecifických prípadoch, napríklad pri rozsvietení reflektorov auta na inak pomerne tmavom videu, môže byť z hľadiska zachovania percepcie nemennosti okolia prínosné posudzovať jasnosť vrámci rozsahu intenzít naprieč celým videom, nie len v rámci daného záberu, avšak z hľadiska rozoznania informácie, čo je pre doménu bezpečnostných kamier kľúčovým, by šlo o degradáciu. Pri prieskume súčasného využívania ekvalizácie histogramu vo videách sme sa však na internete stretli len s individuálnou aplikáciou po snímkach, a to aj pri prístupoch navrhovaných odbornou obcou [30].

Keďže pracujeme primárne s obrazmi reprezentovanými na úrovni farebných kanálov (RGB), a kontrast nutno posudzovať vzhľadom na globálne hladiny intenzity jednotlivých pixelov, teda uniformne, bez ohľadu na farebný kanál, použijeme formát, ktorý oddeľuje komponent jasnosti do separátneho kanálu, na ktorý budeme ekvalizáciu histogramu aplikovať, zatiaľ čo farbu uschováme v kanáloch, ktoré nezmeníme. Takýmto formátom je napríklad YCbCr, využívajúci zložku Y na reprezentáciu jasu, a zložky Cb a Cr na reprezentáciu modrého a červeného chrominančného komponentu.

Praktickým nedostatkom základnej ekvalizácie histogramu v mnohých scenároch sú prípady, kedy pri redistribúcii síce zvýrazníme jasové rozdiely vo väčšine obrazu, no zlejeme ich v nejakej podčasti, ktorá pritom pre nás nesie veľký význam z hľadiska poskytnutej užitočnej informácie. Tento jav by sa mohol manifestovať napríklad presvetlením tváre nasvietenej umelým zdrojom svetla, kedy by klasická ekvalizácia histogramu využila väčšinu z rozsahu zobraziteľných jasov na reprezentáciu tmavého okolia. Preto využijeme aj metódu adaptívnej ekvalizácie histogramu, ktorá posudzuje záber po blokoch (disjunktných, a to rádovo o uhlopriečke v spodných desiatkach pixelov, prípadne jednotkách pixelov) a ekvalizuje pre ne histogramy nezávisle. Aby pri tomto nedochádzalo k zvýrazneniu šumu, pre histogram každého bloku platí limit kontrastu, ktorý ak nejaké pixely prekročia, sú ich intenzity pred samotnou ekvalizáciou orezané, a prevyšujúce hodnoty intenzity uniformne prerozdelené medzi pixely prislúchajúce k iným intenzitám.

2.4.4 Odstránenie rozmazania

Odstránenie rozmazania spôsobeného pohybom objektov v scéne alebo pohybom samotnej kamery, ak ide o takú inštaláciu, síce nemá viditeľný význam z hľadiska sledovaného videa, keďže ľudský mozog si informácie z jednotlivých záberov spája do súvisle definovaného obrazu, ale má význam z hľadiska individuálneho záberu, ktorý môže byť niekedy nutným k identifikácii pohybujúcej sa osoby či vozidla, najmä pri

retrospektívnom vyšetrovaní, a tiež má význam v rámci predspracovania pre ďlašie algoritmy, obzvlášť pre neurónové siete.

Hoci existujú aj neurónové siete trénované špecificky na úlohu odstraňovania rozmazania fotografie, my si vyberieme matematickú metódu, a to dekonvolúciu vo frekvenčnom spektre, pričom budeme predpokladať, že k rozmazaniu došlo aplikáciou gaussovského vyhladzovacieho filtra, a pokúsime sa toto vyhladenie zvrátiť a dostať tak obraz bez rozmazania. Nasadenie neurónových sietí na túto marginálnu úlohu, kde čo najdeterministickejšie dopočítanie informácie preferujeme pred jej odhadom, by bolo v našej doméne použitia nadbytočným, až kontraproduktívnym krokom, z hľadiska časovej náročnosti a nespoľahlivosti odhadu unikátnych čŕt v obraze. Hoci nad postupnosťou záberov tvoriacich video by malo byť možné vytvoriť informovanejší algoritmus odstraňujúci rozmazanie jednotlivých snímok, v čase písania tejto práce neregistrujeme žiaden vhodný pre našu aplikáciu. Viaceré algoritmy sa venujú odstraňovaniu rozmazania spôsobeného pohybom človeka natáčajúceho video kamerou držanou v rukách, avšak tento problém nie je dostatočne podobný nášmu.

2.4.5 Zvýšenie rozlíšenia

OpenCV poskytuje podporu pre zvyšovanie rozlíšenia ako matematickými metódami, tak za pomoci hlbokých neurónových sietí. Z matematických metód budeme využívať bikubickú interpoláciu. Pre využitie neurónových sietí na zvyšovanie rozlíšenia obrazu disponuje OpenCV rozhraním dnn_superres so štyrmi rôznymi algoritmami. Modely ku všetkým vieme stiahnuť z ich verejných githubových repozitárov. Tými sú:

- EDSR (Enhnanced Deep Super-Resolution) [31]
- ESPCN (Efficient Sub-Pixel Convolutional Network) [32]
- FSRCNN (Fast Super-Resolution Convolutional Neural Network) [33]
- LapSRN (Laplacian Pyramid Super-Resolution Network) [34]

Aby sme názorne demonštrovali schopnosti jednotlivých metód zvyšovania rozlíšenia, budeme sa sústrediť na výrez z videa dostatočne malý na to, aby sme po jeho zväčšení daným testovaným násobiteľom nepresiahli rozmery obrazovky, a nemuseli obraz opätovne preškálovávať, čím by sme kompromitovali výpovednú hodnotu porovnania týchto algoritmov.

2.4.6 Rekolorizácia

Veľká časť bezpečnostných kamier si už dnes pomáha infračerveným spektrom (760 nm – 1 mm), alebo jeho časťou stredných až dlhších vlnových dĺžok označovanou ako termovízne spektrum (8–15 μm). Zatiaľ čo viditeľného svetla je v neosvetlených, alebo slabo

osvetlených nočných priestoroch tak málo, že na jeho zachytávanie v čase behu videa by sme potrebovali minimálne kameru s mimoriadne rozmerným objektívom (ktorá by bola drahá a v neposlednom rade nápadná), vo vlnových dĺžkach infračerveného spektra vidíme výraznejšie práve to, čo nás pri dohľadových systémoch zaujíma zďaleka najviac. To sú objekty teplejšie od ich okolia – predovšetkým ľudské postavy – a ich pohyb. Pri tejto práci sme sa stretli s dvomi hlavnými spôsobmi využitia infračerveného spektra.

- S paralelným záznamom v RGB aj v infračervenom spektre, kedy máme oba záznamy k dispozícii samostatne, a žiadne informácie z jedného nie sú na žiadnej vrstve spracovania a zobrazenia prenášané do druhého. Na túto aplikáciu a možnosť kombinácie vizuálnych informácii z oboch spektier a ich zvýraznenie sa pozrieme v ďalšej podkapitole 2.4.7.

- S takzvaným prisvecovaním, kedy kamera automaticky kombinuje termovízny a RGB obraz do šedotónového. Na rekolorizáciu tohto obrazu sa zameriame v tejto podkapitole.

Všetky seriózne metódy rekolorizácie šedotónového obrazu fungujú na základe využitia znalostí [35], keďže neexistuje spôsob, ako správne deterministicky premietnuť hodnoty na spektre medzi bielou a čiernou do farebného priestoru. Tieto znalosti môžu byť dodané asistenciou ľudského používateľa na začiatku behu algoritmu, napríklad formou voľby farebného odtieňa pre vybrané časti obrazu, alebo môžu byť extrahované z nejakej databázy exemplárnych obrazov či vzoriek, alebo môžeme využiť hlboké neurónové siete, ktoré po prvotnom kvalitnom natrénovaní modelu už nepotrebujú prehľadávať množstvá dát v snahe získať referenciu k spracúvanému záberu. Šedotónový obraz, ktorý je výsledkom prisvecovania, sa pokúsime rekolorizovať bez nutnosti ľudskej asistencie, ktorá by v oblasti použitia bezpečnostných kamier mohla byť ľahko vnímaná ako značná nevýhoda systému, keďže by kládla zvýšené nároky na kvalifikovanosť pracovníkov obsluhujúcich dohľadový systém. Napriek existencii algoritmov využívajúcich pre referenciu obraz z tej istej kamery zhotovený za denného svetla [4], [36], takéto metódy zatiaľ nevieme otestovať, pretože k nim nie je poskytnutý zverejnený kód, a pokúšať sa o implementáciu experimentálnych konceptov nie je v možnostiach rozsahu tejto práce, ktorá cieli na vyššiu kvantitu ďalších metrík, o ktorých zlepšenie sa snažíme. Na kolorizáciu preto využijeme hlbokú neurónovú sieť z projektu DeOldify. Napriek tomu, že zverejnený kód je zamýšľaný primárne na použitie s prostredím Jupyter Notebook, ktoré využívajú aj všetky ukážky, manuály a dostupná dokumentácia na internete, pokúsime sa releventné metódy tohto projektu zapracovať priamo do vyvíjaného programu, tak ako aj ostatné z algoritmov, ktoré plánujeme zahrnúť do testovania. Výhoda projektu DeOldify oproti iným je, že má aj metódu špecializovanú na

rekolorizáciu videa, nie len jednotlivých záberov individuálne, čo by bolo jednak menej efektívne a tiež by vzhľadom na neurčitosť problému mohlo viesť k nekonzistencii.

2.4.7 Detekcia pohybu pomocou IR/Termo spektra

Pri kamerových systémoch, ktoré vyhotovujú a paralelne ukladajú záznam aj v RGB spektre, aj v infračervenom (termovíznom) spektre, vieme toto spektrum vlnových dĺžok väčších od viditeľného svetla využiť na obohatenie RGB záberov o informáciu. Túto informáciu ale môžeme rovno posunúť do kontextuálnej domény a využiť ju na detekciu pohybujúcich sa objektov. To by síce bolo možné aj vo farebnom spektre, napríklad subtrakciou pozadia, ktoré je pri stacionárnych bezpečnostných kamerách (čo je väčšina) statické, ale len dokým by sme mali sledovaný priestor dostatočne osvetlený. V tme by schopnosť detekcie výrazne upadala - pohyb by bolo ťažké odfiltrovať od šumu, ak by vôbec bol viditeľný. Vďaka teplote vyššej oproti svojmu okoliu (ak odmyslíme tropické noci v špecifických končinách, no aj tam by sme dokázali vidieť kontrast človeka voči okoliu vyhriatému na akúkoľvek teplotu, keďže zachytené tepelné žiarenie pohybujúcej sa postavy je rôzne v závislosti od termálnej izolácie oblečenia na rôznych častiach tela, a obvykle neprekrytia tváre), však dokážeme detekovať pohyb ľudí v IR spektre za akýchkoľvek svetelných podmienok.

Keďže rozlíšenie videa z objektívu snímajúceho v infračervenom spektre je obvykle nižšie, než z objektívu snímajúceho vo farbe a často nepokrýva ani jeho celé zorné pole, našou úlohou bude najprv spárovať a naškálovať infračervený, respetktíve v prípade našej kamery Annke T200 termovízny obraz, s výsekom z RGB obrazu. V rámci tejto práce si zvolíme jednoduchú a spoľahlivú metódu, kedy na úvod jednorazovo požiadame používateľa, aby v okne, na ktorom je po jednej snímke z každej kamery z rovnakého času, vyznačil najprv ľubovoľný bod v termovíznom obraze (taký, aký ľahko jednoznačne označí aj v druhom spektre), potom tento istý bod vyznačí v RGB obraze, a to isté spraví pre ešte jeden iný ľubovoľný bod. Náš program automaticky nájde transformáciu, ktorú potom aplikuje na informáciu v termovíznom obraze, aby ju zobrazil v RGB obraze na rovnakom mieste relatívne k objektom v scéne. Touto informáciou budú kontúry objektov, extrahovaných subtrakciou pozadia, teda odčítaním po sebe idúcich záberov v binárnom formáte, ktoré nám umožnia vidieť všetky pohybujúce sa objekty s farbou odlišnou od ich pozadia (teda predovšetkým ľudí).

Kapitola 3

Výsledky a vyhodnotenie práce

V tejto kapitole si porovnáme účinok implementovaných metód a ich kombinácií na testovacie videá a špecifiká scenárov, ktoré tieto videá zachytávajú. Taktiež porovnáme ich časovú náročnosť. Hoci sa budeme zmieňovať aj o celkovom čase behu algoritmu alebo ich kombinácii nad celými konkrétnymi videami, tak ako sú aj tieto algoritmy stopované, od začiatku jeho aplikácie na prvý záber po jej dokončenie na poslednom, nezarátavajúc čas potrebný na okolité všeobecnejšie pracovné úkony vykonávané programom, hlavnou jednotkou časovej náročnosti by z intuície mal byť čas potrebný na spracovanie jednoho záberu, a teda celkový čas behu algoritmu vydelený počtom záberov. Avšak nesmieme zabúdať, že na čas behu algoritmu vplývajú aj rozmery samotných záberov, keďže pratkicky všetky z operácií sú aspoň v niektorej zo svojich fáz vykonávané po individuálnych pixeloch alebo ich zoskupeniach. Rolu hrá aj množstvo ďalších faktorov, no tie už sú špecifické pre scénu a informáciu individuálnych videí a nedajú sa parametrizovať tak, ako počet záberov a počet pixelov, a tiež je ich vplyv na čas obvykle menší v porovnaní s časovými a priestorovými atribútmi videí. Preto si ako jednotku časovej náročnosti algorimtu zavedieme mikrosekundu na pixel (μs/px).

V predošlej kapitole 2 sme si predstavili zozbierané videá na ktorých budeme metódy testovať, aj kategórie týchto metód podľa toho, na aký aspekt zvýraznenia informácie pozorovateľovi alebo ďalšiemu algoritmu pri predspracovaní mierime. Keďže väčšinou najvyššiu mieru celkového zlepšenia dosiahneme kombinovaným nasadením viacerych metód, kde často bude aj značne záležať na ich poradí, nemá veľmi význam postupovať pri prezentovaní výsledkov po jednotlivých metódach tak, ako boli predstavené v predošlej podkapitole 2.4. Postupovať po jednotlivých videách by síce nepostrádalo význam, avšak mohlo by vniesť dojem špecifickosti a neuniverzálnosti, a bolo by obtiažnejšie sledovať trendy a opakujúce sa situácie a efekty nezávisle od konkrétnych scén. Budeme teda postupovať po akýchsi scenároch. Príkladom scenáru je napríklad stredne vzdialený človek, pohybujúci sa v tme, alebo scéna s vysokým dynamickým rozsahom. Jednoducho, určité výzvy, s ktorými sa bezpečnostné kamery môžu stretnúť

v reálnom nasadení (samozrejme nie nutne všetky, ale väčšia či menšia podmnožina). Do týchto scenárov potom budú reprezentatívne spadať testovacie videá.

Podstatným vyzrozumením je, že väčšina kombinácii algoritmov vo veľa prípadoch v skutočnosti nevedie k zlepšeniu a môže práveže viesť k degradácii obrazu, najčastejšie pridaním artefaktov. Ide teda o starostlivú voľbu správnych mechanizmov, so správnymi parametrami a v správnom poradí aplikácie podľa toho, čo je našim cieľom, a teda potenciálne, čo sme v jeho prospech ochotní obetovať - napríklad realistickosť obrazu. Mnohé zo zlepšení, ktoré sme dosiahli, by sa nemuseli javiť ako zlepšenia z hľadiska umeleckého dojmu a odrážania reality čo najbližšej percepcii ľudským zrakom, aké sa snažíme dosiahnuť pri natáčaní videí na vlastné účely alebo v kinematografii. Ide o zlepšenia cieliace na zvýraznenie informácie pre účely identifikácie osôb, objektov a činností v sledovanom priestore. To je totiž úloha bezpečnostných kamier. Netúžime po realistickej tme ani plynulosti obrazu na úkor informácie o obsahu priestoru. Neznamená to ale ani to, že úplne ignorujeme rušivé efekty aplikácie algoritmov, keďže ich koncovým respondentom je najčastejšie človek, ktorý musí z obrazu dané informácie komfortne prijímať, a premietnuť si ich do realistického mentálneho obrazu kontextu scény. V niektorých prípadoch nakoniec dôjdeme k zisteniu, najmä za dobrých svetelných podmienok a pri kvalitnej kamerovej aparatúre, že najlepší obraz je z pôvodného videa, a akékoľvek z uvažovaných algoritmov sú kontraproduktívne. V zmysluplne veľkej časti prípadov je to však opačne, a došlo aspoň k malému zlepšeniu kvality záberov, ktoré môže byť rozhodujúce pri potvrdení či vyvrátení dopytovaných skutočností zdokumentovaných videom. Pozrime sa teda na konkrétne ukážky...

3.1 Identifikácia osoby za dobrých svetelných podmienok

Za denného svetla, prípadne v kompletne umelo osvetlenom priestore, obvykle zábery scény neobsahujú miesta, ktoré by vyžadovali algoritmické pozdvihnutie informácie. Kontrast medzi jednotlivými objektmi a črtami je dostatočný, ako vidíme na obrázku 3.1, kde sme porovnávali originálne video s ekvalizáciou histogramu a adaptívnou ekvalizáciou histogramu v dvoch rôznych nastaveniach, a síce došlo k prisvieteniu tienistých miest, ako napríklad pri stene prízemia budovy napravo, ale je otázne, či toto zlepšenie bolo nutné a stojí nám za zhoršenie celkovej prehľadnosti neprirodzene pôsobiacich záberov. Taktiež na tvári autora tejto práce v popredí záberu je vidno, že zvýšenie kontrastu neprinieslo obohatenie vo veci vnímaného detailu, keďže ten už bol dostatočný. Demonštrovali sme však, že dokážeme eliminovať vplyv tieňov v scéne bez kompromitácie informácie, takže pre isté prostredia môže byť ekvalizácia histogramu prínosná.

K zvýrazneniu detailu môže prispieť zaostrovanie, ako môžeme vidieť na obrázku

Obr. 3.1: Vľavo hore: originálny záber, vpravo hore: ekvalizácia histogramu (v YCbCr formáte), vľavo dole: adaptívna ekvalizácia histogramu s okolím 8 pixelov a jeho limitom kontrastu 40, vpravo dole: adaptívna ekvalizácia histogramu s okolím 4 pixelov a jeho limitom kontrastu 20

3.2. Za jemnejšiu formu zaostrovania môžeme považovať aplikáciu odčítania obrazu rozmazaného gaussovským filtrom, v našom prípade berieme 1.5 násobok hodnoty pixelov v pôvodnom obraze a odpočítame 0.5 násobok hodnoty v rozmazanom obraze s gaussovským filtrom rozmerov 3 × 3. Agresívnejší účinok, ak je preferovaný, dosiahneme aplikáciou hranového filtra, v našom prípade v podobe [[0, -1, 0], [-1, 5, -1], [0, -1, 0]]. Na potlačenie rušivého prílišného zvýraznenia štruktúr v okolí bez znehodnotenia unikátneho detailu (ako na ľudských tvárach, je efektívne pred zaostrením hranovým filtrom najprv aplikovať odstránenie šumu nelokálnym priemerovaním. Toto sa javí ako najlepšia možnosť zvýraznenia detailov za denného svetla, avšak jej nevýhodou je, že takéto "bezstratové"(vzhľadom na podstatný detail) odstraňovanie šumu je pomalý proces, a teda nie je aplikovateľné v reálnom čase, len retrospektívne. Kým aplikácia hranového filtra na jeden záber o veľkosti 100 × 100 pixelov (výsek, no škálovanie je lineárne) zabrala len približne 150-200 μs, čiže 0,015 - 0,02 μs/px, a zaostrovanie odčítaním rozmazaného obrazu (tiež s lineárnym škálovaním) 200-300 μs, čiže 0,02 - 0,03 μs/px, takže nie je problém ich aplikovať pri frekvencii živého prenosu 30 alebo aj viac snímok za sekundu, odstraňovanie šumu nelokálnym priemerovaním si vyžiadalo zhruba 345000 μs aj na takomto malom výseku, čiže 34 μs/px. Je možné ho bez citeľnej straty kvality zrýchliť znížením počtu okolitých záberov, v ktorých hľadáme referenčné výseky pre priemerovanie (tu sme využívali 2 zábery pred a 2 po), no toto zrýchlenie

zodpovedá počtu prehľadávaných záberov, a teda nedokážeme dosiahnuť hodnotu pod 6 μs/px.

Obr. 3.2: Vľavo hore: originálny záber, vpravo hore: zaostrenie odčítaním rozmazaného obrazu, vľavo dole: zaostrenie hranovým filtrom, vpravo dole: odstránenie šumu a následné zaostrenie hranovým filtrom

Ďalšie zlepšenie detailu dosiahneme použitím predtrénovaných neurónových sietí na zvýšenie rozlíšenia. Porovnaniu implementovaných modelov sa budeme venovať v ďalších podkapitolách, no ukážku práce najkvalitnejšej, no aj časovo najnáročnejšej z nich - EDSR, môžeme vidieť na obrázkoch 3.3, kde kombinujeme zaostrovanie a EDSR, a 3.4, kde ešte pridávame ekvalizácie histogramu. Je nutné podotknúť, že narozdiel od predchádzajúcej ukážky, originálny záber tu nie je bikubicky interpólovaný (je interpólovaný metódou najbližšieho suseda), aby sme zdôraznili dotvorenie pixelov EDSR pri 4-násobnom zvýšení rozlíšenia. Za kombinácie metód s najpriaznivejším efektom

môžeme označiť:

- zaostrenie hranovým filtrom a následné zvýšenie rolíšenia EDSR
- zaostrenie odčítaním rozmazaného obrazu, následné zvýšenie rolíšenia EDSR a ekvalizácia histogramu na záver

Rýchlosť nasadenia neurónových sietí bude analyzovaná ďalej v práci, je však nutné zdôrazniť, že žiaden z implementovaných modelov nie je dostatočne rýchly na nasadenie v reálnom čase, naopak všetky implementované kombinácie zaostrovania a ekvalizácii histogramov sú, minimálne pre videá do veľkosti 1280 × 720 pixelov.

3.2 Identifikácia osoby za nepriaznivých svetelných podmienok

Identifikáciu osoby v tme sme testovali:

- Bez akejkoľvek pomoci v spektre vlnovej dĺžky mimo viditeľného svetla.
- Na záberoch z kamery Hikvision s infračerveným prisvecovaním.
- S využitím termovízie k detekcii pohybujúcej sa osoby.

Na videu Annke_RGB_dark sme otestovali 34 kombinácii algoritmov, hlavne ale na výreze o veľkosti 400 × 200 pixelov, zachytávajúcom autora práce v tmavej časti scény preskakujúceho zábradlie a otočiaceho sa tvárou do kamery. Pri ekvalizácii histogramu, sme narazili na problém zmienený v podkapitole 2.4, kedy tvár osoby (a na videu k tomu biele tričko) predstavujú len malý zlomok pixelov scény, no zároveň sú najsvetlejšími časťami scény. Ekvalizácia histogramu ich pixely potom zleje do menšieho rozpätia intenzít, aby dosiahla zlepšenie separácie intenzít v tmavej väčšine scény. K eliminácii tohto problému by mala viesť adaptívna ekvalizácia histogramu, ktorá zvyšuje kontrast na lokálnej báze. Problémom pri objektoch v takto nízkom rozlíšení sú veľmi silné artefakty, ktoré ňou vziknú a rozbijú akúsi integritu a plynulosť obrazu. Ani jeden z implementovaných spôsobov ekvalizácií histogramu tak nie je vhodný v častom scenári, kedy sa o čosi svetlejšia osoba pohybuje v tmavom priestore v strednej vzdialenosti od kamery, ako je vidieť na ukážke 3.5, kde sme ekvalizácie nasadili po odstránení šumu a v kombinácii s oboma metódami zaostrovania.

Pri zvyšovaní rozlíšenia tohto výrezu si porovnáme schopnosti a časovú náročnosť všetkých implementovaných modelov neurónových sietí: EDSR, ESPCN, FSRCNN a LapSRN, a to pri 2-násobnom zväčšení (na inom videu si ich porovnáme aj pri 4-násobnom). Ako vidíme na zábere 3.6, rozdiely sú len ťažko badateľné, minimálne a v podstate zanedbateľné.

Obr. 3.3: Vľavo hore: originálny záber v pôvodnom rozlíšení, vpravo hore: 4-násobné zvýšenie rolíšenia EDSR, vľavo dole: zaostrenie hranovým filtrom a následné 4-násobné zvýšenie rolíšenia EDSR, vpravo dole: zaostrenie odčítaním rozmazaného obrazu a následné 4-násobné zvýšenie rolíšenia EDSR

Zanedbateľnými ale nie sú rozdiely v časovej náročnosti použitia jednotlivých modelov - naopak - sú značné. Nasadenie EDSR - modelu, ktorý sa nám vo väčšine aplikácií ukázal ako najschopnejší, no len s minimálnym, často zanedbateľným, no stále niekedy potenciálne rozhodujúcim rozdielom - trvalo len na spomínaný výsek o rozmeroch 400 $\times$ 200 pixelov v priemere 18835037 μs na každú snímku, teda 235 μs/px. Použitie tohto modelu je možné len retrospektívne a pri opodstatnených prípadoch. Pre predstavu, na zvýšenie rozlíšenia 135 snímok tohto podobrazu, teda 4 a pol sekundy videa, algoritmus potreboval 2543 sekúnd. Na rovnakom úseku videa to zabralo modelu ESPCN 8,85s, teda 65555 μs na záber a 0,82 μs/px. V prípade FSRCNN to bolo 19,57s pre

Obr. 3.4: Vľavo hore: originálny záber v pôvodnom rozlíšení, vpravo hore: zaostrenie hranovým filtrom, následné 4-násobné zvýšenie rolíšenia EDSR a ekvalizácia histogramu na záver, vľavo dole: zaostrenie odčítaním rozmazaného obrazu, následné 4-násobné zvýšenie rolíšenia EDSR a ekvalizácia histogramu na záver, vpravo dole: zaostrenie odčítaním rozmazaného obrazu, následné 4-násobné zvýšenie rolíšenia EDSR a adaptívna ekvalizácia histogramu na záver s okolím 4 pixelov a jeho limitom kontrastu 20

Obr. 3.5: Vľavo hore: odstránenie šumu, následné zaostrenie hranovým filtrom a ekvalizácia histogramu na záver, vpravo hore: ostránenie šumu, následné zaostrenie hranovým filtrom a adaptívna ekvalizácia histogramu na záver s okolím 4 pixelov a jeho limitom kontrastu 20, vľavo dole: ostránenie šumu, následné zaostrenie odčítaním rozmazaného obrazu a ekvalizácia histogramu na záver, vpravo dole: ostránenie šumu, následné zaostrenie odčítaním rozmazaného obrazu a adaptívna ekvalizácia histogramu na záver s okolím 4 pixelov a jeho limitom kontrastu 20

Obr. 3.6: Vľavo hore: EDSR, vpravo hore: ESPCN, vľavo dole: FSRCNN, vpravo dole: LapSRN

predspracovanie	EDSR	ESPCN	FSRCNN	LapSRN
žiadne	235.44	0.82	1.81	8.26
zaostrenie hranovým filtrom	210.19	0.66	1.84	7.56
odstránenie šumu a zaostrenie hranovým filtrom	184.99	0.68	1.59	7.52
zaostrenie odčítaním rozmazaného obrazu	187.21	0.66	1.68	7.56
odstránenie šumu a zaostrenie odčítaním rozmazaného obrazu	189.09	0.68	1.54	7.58

Tabuľka 3.1: Porovnanie časov implementovaných modelov pri zdvojnásobení rozlíšenia v μs/px

celý výsek videa, čiže 144962 μs na záber a 1,81 μs/px . Použitie LapSRN nás vyšlo na 89,28s celkovo, 661333 μs na záber a 8,26 μs/px. Okrem toho sme zvýšenie rozlíšenia v ďalších testoch kombinovali aj s predspracovaním, kde sme vyskúšali obe implementované metódy zaostrovania obrazu, pričom každú z nich aj samostatne, aj na obraze zbavenom šumu (pred nasadením neurónovej siete). Porovnanie rýchlostí jednotlivých modelov v závislostu od predspracovania sa nachádza v tabuľke 3.1.

Keďže výsledky aplikácie jednotlivých modelov sú si v tomto scenári prakticky na nerozoznanie blízke, budeme porovnávať aplikácie ďalších algoritmov na toto video v kombinácii s najrýchlejším z nich - ESPCN, a druhým najrýchlejším - FSRCNN. Preskúmali sme aj kombinácie s ďalšími, no trendy vzťahujúce sa na výsledky kombinácii s ďalšími metódami sú medzi nimi rovnaké. Ako vidíme na obrázku 3.7, predspracovanie zaostrovaním hranovým filtrom vytvára nezanedbateľné artefakty aj na tvári osoby, a to aj v kombinácii s odšumeným obrazom. Zaostrovanie odčítaním rozmazaného obrazu je teda preferované.

Odstránenie šumu sa tu nepreukázalo dostatočne citeľne prínosné na obhájenie jeho vysokej časovej náročnosti. Aj ak používame k hľadaniu refrenčných výsekov len aktuálny záber, čas potrebný na túto operáciu vyše 23-násobne prevyšuje čas trvania videa, hoci ju aplikujeme len na jeho výrez o veľkosti 400 $\times$ 200 pixelov. Jej náročnosť tu je teda na úrovni nie menšej než 0,7 sekundy na záber, alebo 9,68 μs/px. Preto ju v tomto scenári nebudeme ďalej aplikovať. Opačné poradie aplikácie zaostrovania, kedy obraz zaostríme až po zvýšení rozlíšenia, vzniknuté artefakty do solídnej miery eliminuje, ako je vidieť na ukážke 3.8, a prezentuje nám tak ponuku najlepších postupností zlepšení, aké dokážeme na tomto videu pomocou testovaných algoritmov dosiahnuť.

Ďalšie videá spadajúce do scenára osoby pohybujúcej sa v tme, sú zaobstarané kamerou Hikvision s infračerveným prisvecovaním a originálne sú v šedotónovej podobe. Našim prvým krokom teda bude ich rekolorizácia pomocou neurónovej siete z projektu

Obr. 3.7: Vľavo hore: zaostrenie hranovým filtrom + ESPCN, vpravo hore: odstránenie šumu + zaostrenie hranovým filtrom + ESPCN, vľavo dole: zaostrenie odčítaním rozmazaného obrazu + ESPCN, vpravo dole: odstránenie šumu + zaostrenie odčítaním rozmazaného obrazu + ESPCN

Obr. 3.8: Vľavo hore: zaostrenie hranovým filtrom + FSRCNN, vpravo hore: zaostrenie odčítaním rozmazaného obrazu + FSRCNN, vľavo dole: FSRCNN + zaostrenie hranovým filtrom, vpravo dole: FSRCNN + zaostrenie odčítaním rozmazaného obrazu

Obr. 3.9: Vľavo hore: originálny záznam, vpravo hore: rekolorizácia, vľavo dole: rekolorizácia + zaostrenie hranovým filtrom, vpravo dole: rekolorizácia + zaostrenie hranovým filtrom + ekvalizácia histogramu

Deoldify. Tá sa ukázala ako spoľahlivá a produkuje výsledky s odchýlkou od reálnych farieb dostatočne malou na to, aby nešlo len o uveriteľné, ale aj o realistické dofarbenie konkrétnej scény, no len ak má na záberoch zmysluplné referenčné objekty. Farby pôsobia trochu nenaturálnym dojmom, avšak to je zapríčinené aj tým, že nejde o klasický šedotónový obraz, ale obohatený o informáciu z infračerveného spektra. Na ukážke 3.9 vidíme výsledok aplikácie rekolorizácie, a efekt následnej aplikácie zaostrovania, prípadne aj ekvalizácie histogramu. Prostredie, v akom je záznam vyhotovovaný, obsahujúce ihličnatú zeleň tvoriacu vysokú frekvenciu striedania pixelov vysokej a nízkej intenzity v kombinácii so zdrojmi svetla, je značne nepriaznivé pre aplikáciu ekvalizácie histogramu.

Na výstrih z videa o rozmeroch 300 × 240 pixelov, kde autor práce podišiel bližšie ku kamere, sme následne aplikovali 4-násobné zvýšenie rozlíšenia pomocou všetkých implementovaných modelov neurónových sietí na to zameraných. Vzájomné rozdiely medzi nimi boli opäť len veľmi ťažko badateľné, ak vôbec. Toto vidieť na ukážke 3.10. Subjektívne sa nám však ako najlepší javil výsledok modelu LapSRN, keďže nevniesol do obrazu badateľné artefakty, a zároveň zobrazuje jednotlivé objekty na scéne ostro

Obr. 3.10: Na rekolorizáciu nadväzujúce - vľavo hore: EDSR, vpravo hore: ESPCN, vľavo dole: FSRCNN, vpravo dole: LapSRN

a kontinuálne, preto pri ďalších snahách o zlepšenie týchto snímok zúžitkujeme tento algoritmus.

V porovnaní 3.11 vidíme, že najlepší výsledok, ostrý, kontinuálny a bez artefaktov, sme dosiahli záverečnou aplikáciou zaostrovania hranovým filtrom na rekolorizované video zväčšené modelom LapSRN (lepší, než pri inej forme alebo poradí aplikácie zaostrovania). Oproti aplikácii zaostrovania hranovým filtrom na video v pôvodnom rozlíšení v ukážke 3.9 je tu vidieť výrazné zlepšnie, keď sú vďaka zväčšeniu rozlíšenia rozostupy medzi hranami väčšie.

Pre názornosť si môžeme ukázať neúspech odstraňovania rozmazania, ktoré malo slúžiť na predspracovanie záberov pre neurónové siete pred zväčšením rozlíšenia. S ním obvykle automaticky kombinujeme ekvalizáciu histogramu, keďže pri procese, ak nemá vniesť do obrazu veľkorozmerné artefakty, dochádza ku komprimácii intenzít jasu do spektra nízkych hodnôt. Napriek tomu, že táto metóda mala prínosný účinok na testovacích rozmazaných záberoch nepochádzajúcich z naších videí, zjavne je funkčná len v úzkom spektre prípadov. Naše rekolorizované video pri aplikácii po jednotlivých

Obr. 3.11: Na rekolorizáciu nadväzujúce - Vľavo hore: zaostrenie hranovým filtrom + LapSRN, vpravo hore: zaostenie odčítaním rozmazaného obrazu + LapSRN, vľavo dole: LapSRN + zaostrenie hranovým filtrom, vpravo dole: LapSRN + zaostrenie odčítaním rozmazaného obrazu

Obr. 3.12: Rekolorizácia + odstránenie rozmazania v záberoch + ekvalizácia histogramu

kanáloch skoro zbavila farby, no hlave ho úplne dezintegruje, ako vidieť na zábere 3.12. Rozmazanie vyskytujúce sa vo videách z bezpečnostných kamier, ktoré bývajú uchytené stabilne a vzdialene od pozorovaných objektov, však nebýva signifikantným problémom, a teda sa ním nebudeme ďalej zaoberať, a ani sme tu neskúšali efekt tohto predspracovania na následné zvyšovanie rozlíšenia neurónovými sieťami.

V priestore bez akéhokoľvek iného než ambientného nočného osvetlenia, mimo zóny pouličných lámp - praktickú aplikáciu si môžeme predstaviť napríklad u zadného vchodu do nejakej prevádzky - sme zistili, že infračervené prisvecovanie kamery Hikvision žiaľ nie je ani z realtívne blízkej vzdialenosti postačujúce, ak by sme chceli identifikovať tvár osoby. Najprv sme sa pokúsili aplikovať rekolorizáciu na pôvodné, aj šumu zbavené video, ktorá trvala v priemere skoro 3 sekundy na záber o veľkosti 640 $\times$ 360 pixelov, konkrétne pri videu bez predspracovania 12,5 μs/px a s predspracovaním odstránením šumu 11,05 μs/px. Keďže video však neposkytuje prakticky žiadne objekty, ktorým by bola pomerne jednoznačne priraditeľná farba, ako boli na videu Hikvision_dark_distant stromy či pouličné osvetlenie, neurónová sieť je tu bezradná a oba pokusy o kolorizáciu sa vyznačujú striedaním farebných odtieňov videa porovnateľ-

Obr. 3.13: Farebne nesprávny efekt rekolorizácie videa bez referenčných objektov. Zaujímavým by mohlo byť, že vždy práve vtedy, keď autor práce na zázname podliezol, alebo inak interagoval s rámovou konštrukciou sušiaka - tu vidieť - záznam náhle prešiel do inak neprítomného modrého sfarbenia.

ným s umiestnením disko gule neďaleko sledovaného priestoru, a nesprávnym odhadom farby oblečenia, viď ukážka 3.13.

Ďalej sa teda radšej pokúsime zlepšiť šedotónové video. Pri tom sa nám neosvedčili ekvalizácie histogramu, osvedčilo sa však zaostrovanie hranovým filtrom, a to buď ako jediná aplikovaná metóda, ak nám nevadí šum v prostredí, majúc tiež značnú výhodu, že ju vieme nasadiť na živý prenos v reálnom čase, alebo po odstránení šumu nelokálnym priemerovaním, ak preferujeme uhladenejší obraz (no stále nechceme prísť o informáciu). Porovnanie s originálom a zaostrovaním odčítaním rozmazaného obrazu je na obrázku 3.14.

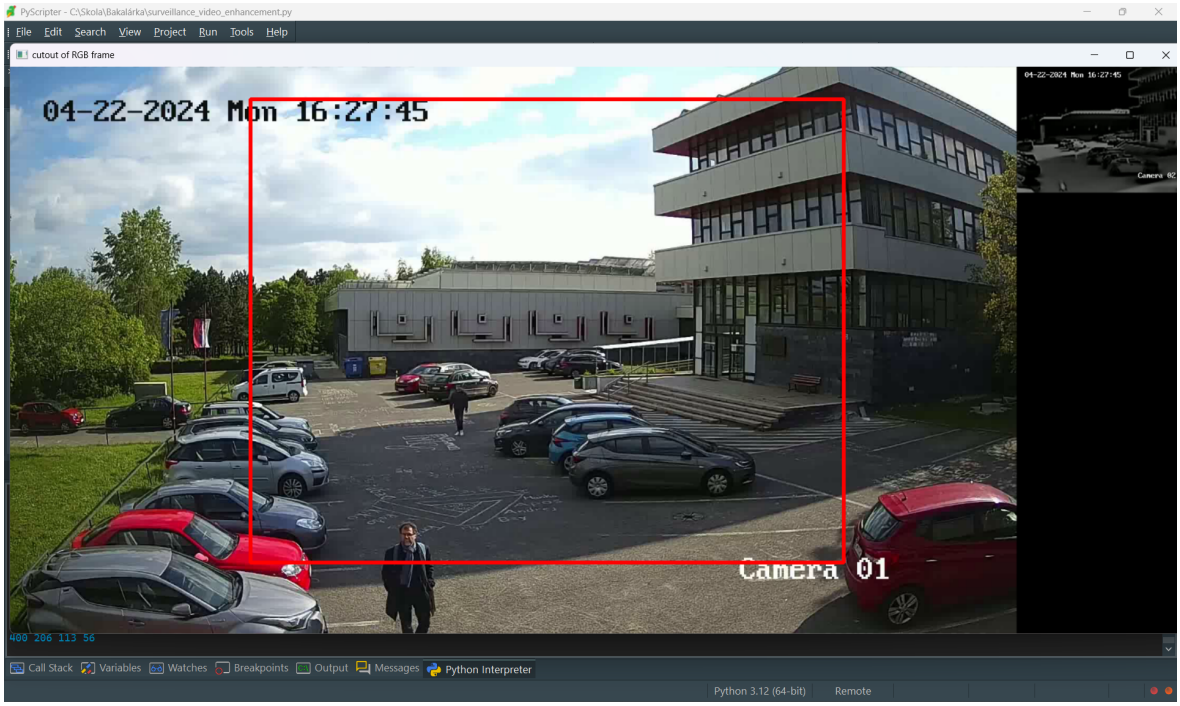
Prvoradé, ak by sme chceli osobu identifikovať, je ju vôbec na zábere zahliadnuť. Ak je napríklad na druhom konci prenosu obrazu z kamery strážnik či vrátnik, mohlo by mu byť nápomcné byť upozornený na pohyb človeka v sledovanom priestore, keby sa tento vyskytol v tmavej časti scény a prípadne by túto skutočnosť využil voľbou tmavého šatstva či dokonca prekrytia tváre. Akákoľvek snaha farebne splynúť s temným pozadím by mu bola zbytočná, ak ho dokážeme vidieť v infračervenom spektre, vďaka teplu, ktoré vyžaruje. Aby ale strážnik nemusel sledovať dve kamery pre jednu scénu zároveň, chceli by sme informáciu o pohybujúcej sa siluete teplejšieho objektu premietnuť na hlavný RGB záznam. Metódu na toto slúžiacu sme implementovali, s tým, že strážnik na začiatku vyznačí dva body reprezentujúce rovnaké miesto na RGB

Obr. 3.14: Vľavo hore: pôvodný záber, vpravo hore: zaostenie odčítaním rozmazaného obrazu, vľavo dole: zaostrenie hranovým filtrom, vpravo dole: odstránenie šumu nelokálnym primerovaním + zaostrenie hranovým filtrom

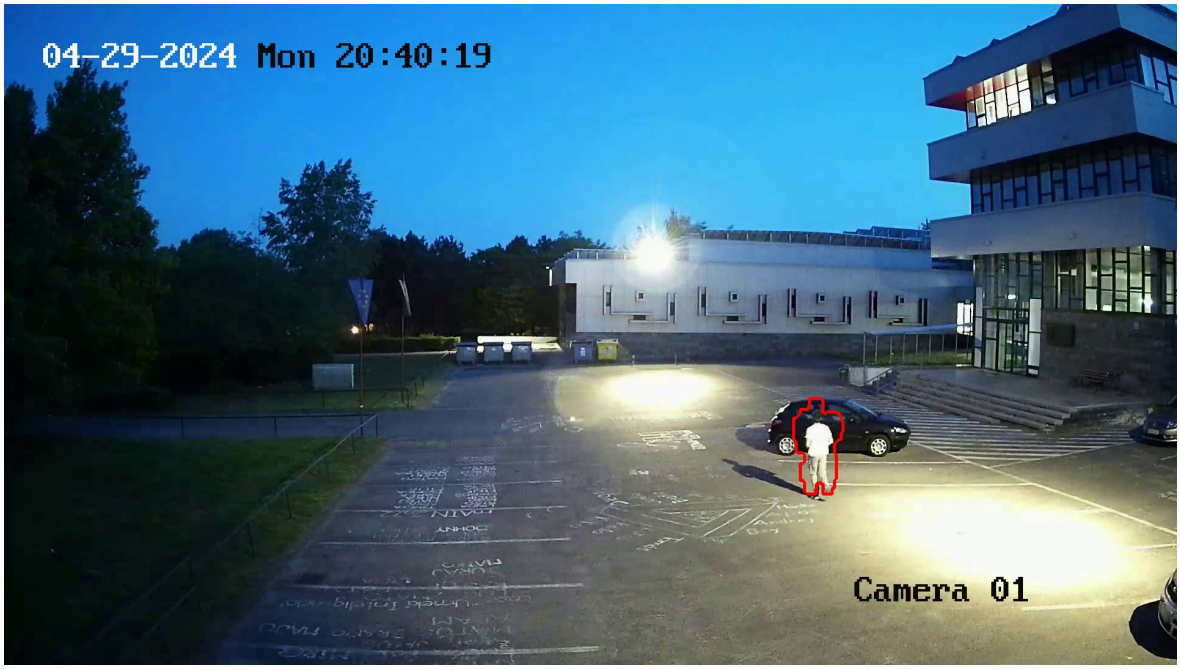
aj infračervenom zábere, a program následne párovanie medzi týmito bodmi využije, aby vypočítal transformáciu medzi scénami oboch spektier, a zvýrazňuje siluetu pohybujúcich sa objektov v RGB spektre na základe informácii z infračerveného spektra. Ukážku napárovaného výseku je možné vidieť na obrázku 3.15, pri testovaní ešte za dňa, kvôli názornosti. Ukážku následného zvýraznenia pohybujúcej sa osoby možno vidieť na obrázku 3.16.

3.3 Identifikácia vozidla

Identifikovať vozidlo sa môžeme snažiť jednoznačne - pomocou evidenčného čísla, alebo typovo - model, ročník, farba, atď - čo je podstatne jednoduchšie. Pre účely porovnania algoritmov sme z videa american_crossroad_traffic spravili výrez špecificky zameraný na autá odchádzajúce od kamery v ich najbližšej pozícii k nej, kde máme najlepšiu šancu identifikovať aj štátnu poznávaciu značku. Rozhodli sme sa tu opäť porovnať predtrénované neurónové siete v záujme zistenia, či sa líšia ich schopnosti v prípade rekonštrukcie textu v nízkom rozlíšení. Toto porovnanie bez akéhokoľvek predspracovania môžeme vidieť na ukážke 3.17. Podľa pohľadu na emblém značky, označenie modelu v pravej hornej časti kufra, a disk kolesa, vidíme, že model EDSR je o čosi schopnejší, než ostatné. Zároveň je ale aj najpomalší. Otázne je, či nám záleží na rýchlosti, ak žiadnu, ani najrýchlejšiu sieť - čiže ESPCN - nedokážeme nasadiť v reálnom čase.



Obr. 3.15: Ukážka napárovania oblasti záberu infračervenej kamery do záberu RGB kamery na základe voľby dvoch referenčných bodov používateľom



Obr. 3.16: Približné označenie siluety pohybujúcej sa osoby získané z infračerveného záberu automaticky vygenerované nami zostrojenou metódou track_movement_thermo_to_rgb.

Obr. 3.17: Vľavo hore: EDSR, vpravo hore: ESPCN, vľavo dole: FSRCNN, vpravo dole: LapSRN

Keďže však časový rozdiel medzi týmito dvoma algoritmami je zhruba 200-násobný, stále považujeme za zaujímavé si ich detailnejšie porovnať v tomto scenári, bohatom na detail.

Ešte raz sme dali pri predspracovaní šancu algorimtu odstraňujúcemu rozmazanie, no skutočne, ako je vidieť na zábere 3.18, ide práveže o veľmi kontraproduktívny proces v našich prípadoch.

Obr. 3.18: Hore: odstránenie rozmazania + ekvalizácia histogramu (nutná po odstraňovaní rozmazania, keďže obraz veľmi stmavne) + EDSR, dole: odstránenie rozmazania + ekvalizácia histogramu + ESPCN

Obr. 3.19: Vľavo hore: zaostrovanie hranovým filtrom + EDSR, vpravo hore: zaostrovanie hranovým filtrom + ESPCN, vľavo dole: zaostrovanie odčítaním rozmazaného obrazu + EDSR, vpravo dole: zaostrovanie odčítaním rozmazaného obrazu + ESPCN

EDSR a ESPCN sme si porovnali nasadené na predspracované obrazy zaostrovaním hranovým filtrom, aj odčítaním rozmazaného obrazu - viď záber 3.19, ako aj s tým, že ešte pred aplikáciou zaostrovania sme aplikovali odstraňovanie šumu nelokálnym priemerovaním - viď záber 3.20. Ak obraz preostríme, EDSR sa ukazuje ako dvojsečná zbraň a vie si domyslieť aj to, čo na zábere nie je, avšak tak, že zdegraduje pôvodnú informáciu. Dalo by sa povedať, že oproti ESPCN má "viac fantázie", avšak často zas vďaka tomu skutočne uhádne detail, ktorý mu štruktúra pixelov obrazu v nižšom rozlíšení napovedá, presnejšie. Výsledky necháme čitateľov posúdiť samých.

Na záver tohto scenáru by bolo vhodné si porovnať zlepšené obrazy s pôvodnými. Porovnáme EDSR aj ESPCN, s oboma metódami zaostrovania, no bez odstraňovania šumu, keďže ide o výpočtovo zdĺhavú operáciu (bezstratovo nelokálnym priemerovaním), ktorá nepriniesla presvedčivé zlepšenie v dobre osvetlenom prostredí. Do porovnania zaradíme aj bikubickú interpoláciu. Ukážky sú na obrázkoch 3.21 a 3.22.

3.4 Prostredia s rozmanitou mierou osvetlenia

Ak je kamera vystavená priamemu slnku, obzvlášť pri jeho východe a západe, dynamický rozsah v scéne predstavuje veľkú výzvu pre bezpečnostné kamery. Veľmi silný prípad takejto situácie máme na zábere dutch_street_sunset. Tu sa nám budú veľmi hodiť ekvalizácie histogramov. Adaptívne ekvalizácie sú efektnejšie, no zanechávajú značné artefakty v podobe akéhosi mozajkovania obrazu. Čím väčší prípustný limit rozdielu kontrastu zvolíme (tento je nutné správne zladiť s veľkosťou okolia - my po-

Obr. 3.20: Vľavo hore: odstránenie šumu nelokálnym priemerovaním + zaostrovanie hranovým filtrom + EDSR, vpravo hore: odstránenie šumu nelokálnym priemerovaním + zaostrovanie hranovým filtrom + ESPCN, vľavo dole: odstránenie šumu nelokálnym priemerovaním + zaostrovanie odčítaním rozmazaného obrazu + EDSR, vpravo dole: odstránenie šumu nelokálnym priemerovaním + zaostrovanie odčítaním rozmazaného obrazu + ESPCN

Obr. 3.21: Vľavo hore: záber v pôvodnom rozlíšení, vpravo hore: štvornásobné zväčšenie bikubickou interpoláciou, vľavo dole: zaostrovanie hranovým filtrom + EDSR, vpravo dole: zaostrovanie odčítaním rozmazaného obrazu + EDSR

Obr. 3.22: Vľavo hore: záber v pôvodnom rozlíšení, vpravo hore: štvornásobné zväčšenie bikubickou interpoláciou, vľavo dole: zaostrovanie hranovým filtrom + ESPCN, vpravo dole: zaostrovanie odčítaním rozmazaného obrazu + ESPCN

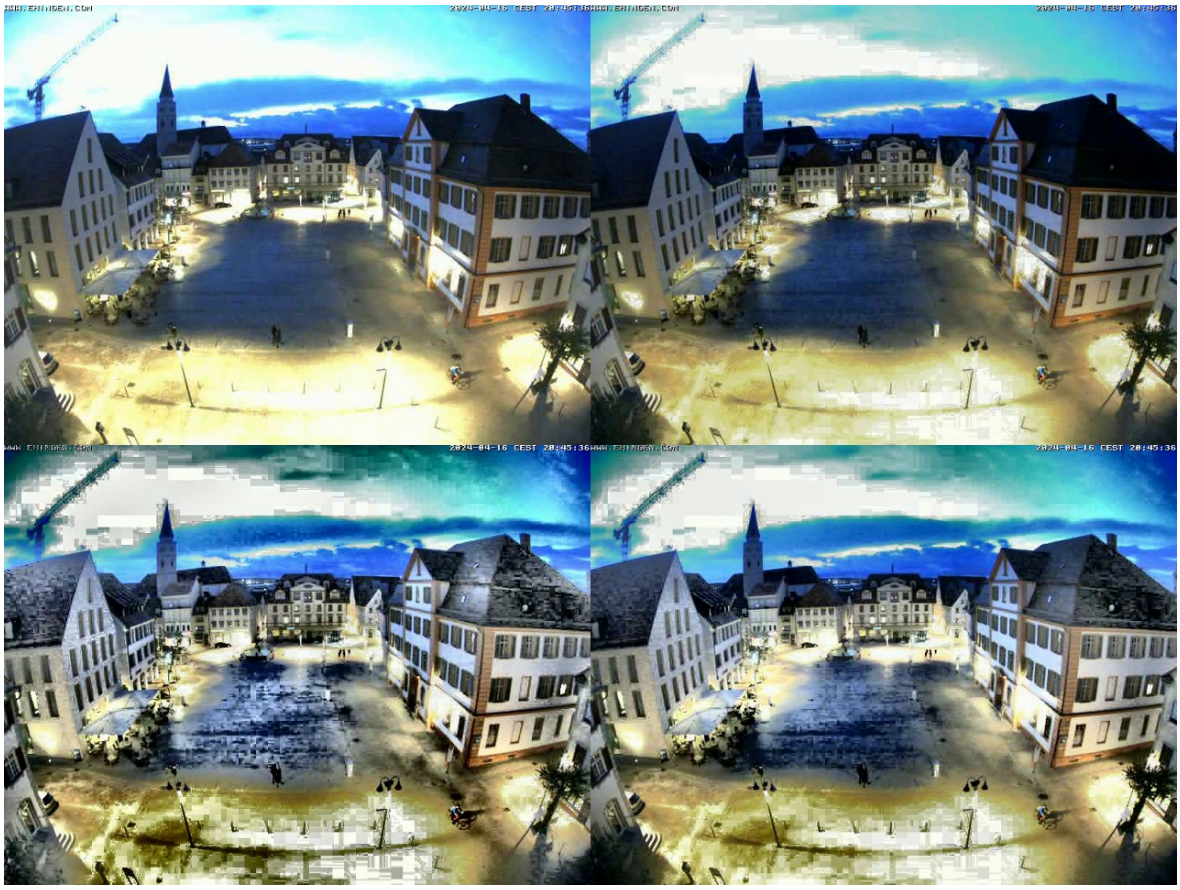
užívame okolie veľkosti 4 s limitom 20 pre hladší efekt, a okolie veľkosti 8 s limitom 40 pre radikálnejšie zvýraznenie), tým lepší kontrast získame medzi inak nevidenými podobrazmi - viď obloha na ukážke 3.23, ktorá demonštruje aj spomínané artefakty, a schopnosti vytiahnuť tiene a kompenzovať presvetlenie v obraze. Opäť tu platí, že čím výraznejšie zlepšenie čitateľnosti informácií v obraze si prajeme, tým menej naturálne bude záber pôsobiť. V tomto prípade si však myslíme, že to určite stojí za to (pri bezpečnostných kamerách), keďže na pôvodnom videu vieme ledva rozlíšiť povrch a tvar cesty od okolia, a len ťažko by sme tam chceli spozorovať nejakú osobu.

Vzhľadom na dostatočne prirodzený vzhľad obrazu bez artefaktov, ktorý zároveň poskytuje vizuálnu informáciu o dianí v ulici, sme si zvolili klasickú globálnu ekvalizáciu histogramu (aplikovanú vo formáte YCrCb), a osvedčilo sa nám ju nasadiť na odšumený obraz (nelokálnym priemerovaním) a kombinovať ju s niektorou z metód zaostrovania. Zaostrovanie hranovým filtrom je pri tom agresívnejšie, kým zaostrovanie odčítaním rozmazaného obrazu tu produkuje len ťažko rozoznateľné zlepšenie, avšak zachováva hladký charakter obrazu, viď ukážka 3.24.

Pozrime sa teraz na trochu miernejšiu situáciu, kde stále čelíme oblohe presvetlenej zapadajúcim slnkom, avšak nie slnku samotnému. Kamera v scéne german_marketplace_eve je síce vzdialenejšia od objektov, ktoré by sme mohli chcieť identifikovať, než dáva zmysel pre bezpečnostné kamery, avšak na zázname je veľa prvkov, na zlepšenie detailu ktorých sa môžeme zamerať. Začali sme s ekvalizáciami histogramu, v záujme vytiahnutia štruktúr na tmavých miestach na zábere. V prípade tohto záznamu, kedy je obraz zašumenejší, než napríklad v predošlom videu dutch_street_sunset, hoci - čo je

Obr. 3.23: Vľavo hore: pôvodný záber, vpravo hore: ekvalizácia histogramu, vľavo dole: adaptívna ekvalizácia histogramu s okolím 8 pixelov a jeho limitom kontrastu 40, vpravo dole: adaptívna ekvalizácia histogramu s okolím 4 pixelov a jeho limitom kontrastu 20

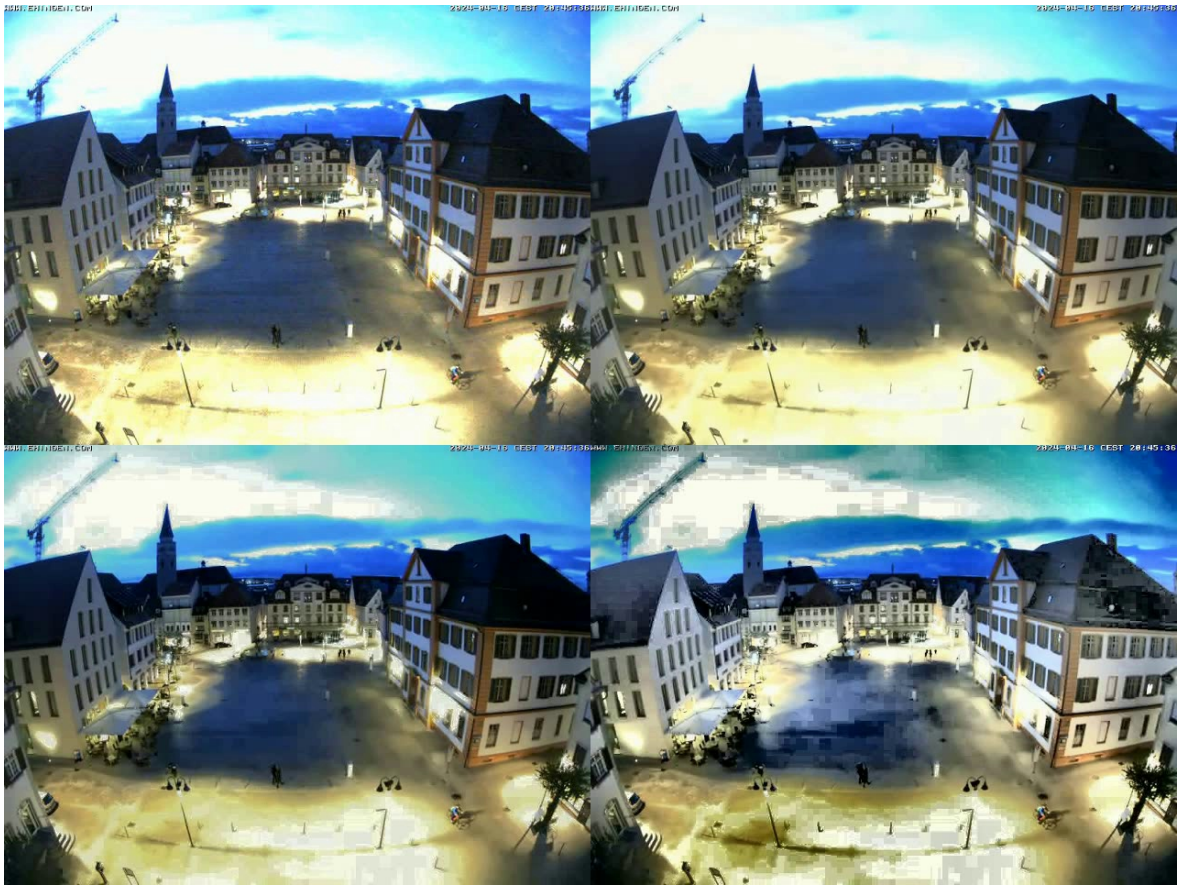
Obr. 3.24: Vľavo hore: pôvodný záber, vpravo hore: odstránenie šumu nelokálnym priemerovaním + ekvalizácia histogramu, vľavo dole: odstránenie šumu nelokálnym priemerovaním + ekvalizácia histogramu + zaostrenie hranovým filtrom , vpravo dole: dstránenie šumu nelokálnym priemerovaním + ekvalizácia histogramu + zaostrenie odčítaním rozmazaného obrazu



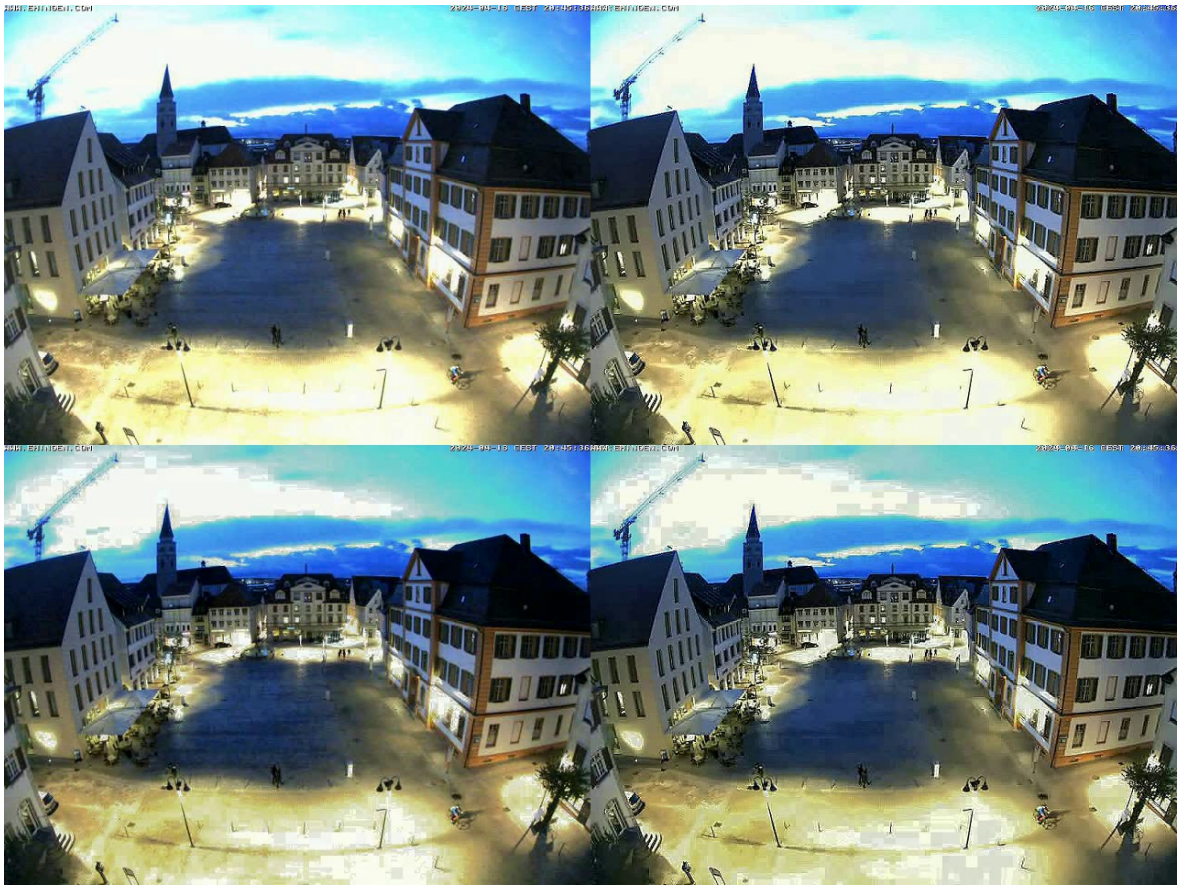
Obr. 3.25: Vľavo hore: pôvodný záber, vpravo hore: ekvalizácia histogramu, vľavo dole: adaptívna ekvalizácia histogramu s okolím 8 pixelov a jeho limitom kontrastu 40, vpravo dole: adaptívna ekvalizácia histogramu s okolím 4 pixelov a jeho limitom kontrastu 20

dôležité uvedomenie - sa tak na prvý pohľad vôbec nemusí zdať, sa veľmi osvedčilo pre potlačenie artefaktov v dôsledku ekvalizácii histogramov, nasadiť najprv na obraz odstraňovanie šumu nelokálnym priemerovaním. Pozitívny dopad tohto predspracovania je vidieť porovnaním medzi ukážkou 3.25 a ukážkou a 3.26, na ktorej síce došlo pri odstránení šumu k miernej degradácii detailu opakujúcich sa štruktúr, avšak z princípu fungovania algoritmu nelokálneho priemerovania ostali unikátne detaily zachované, a to je to, čo nás pri bezpečnostných kamerách zaujíma - detaily jedinečných postáv, vozidiel, prípadne iných objektov, a sme ochotní obetovať detail štruktúry častí budovy, alebo, na našom zázname, napríklad žeriavu.

Opäť sa nám osvedčilo tieto algoritmy kombinovať s niektorou z implementovaných metód zaostrovania. Tiež sme odpozorovali, že je prínosnejšie aplikovať zaostrovanie v poradí ešte pred ekvalizáciou histogramu. Kvôli prítomnosti artefaktov pri adaptívnej ekvalizácii histogramu považujeme aj tu globálnu ekvalizáciu v YCrCb formáte za vhodnejšiu. Jej najúspešnejšie kombinácie s algoritmami zaostrovania vidíme na zábere 3.27. Čitateľovi odporúčame zamerať sa na cyklistu vpravo dole a stánky reštaurácii naľavo. Ak chceme zlepšenie aplikovať v reálnom čase, nemôžeme si dovoliť zbavova-



Obr. 3.26: Vľavo hore: pôvodný záber, vpravo hore: odstránenie šumu nelokálnym priemerovaním, vľavo dole: odstrénenie šumu nelokálnym priemerovaním + ekvalizácia histogramu, vpravo dole: odstrénenie šumu nelokálnym priemerovaním + adaptívna ekvalizácia histogramu

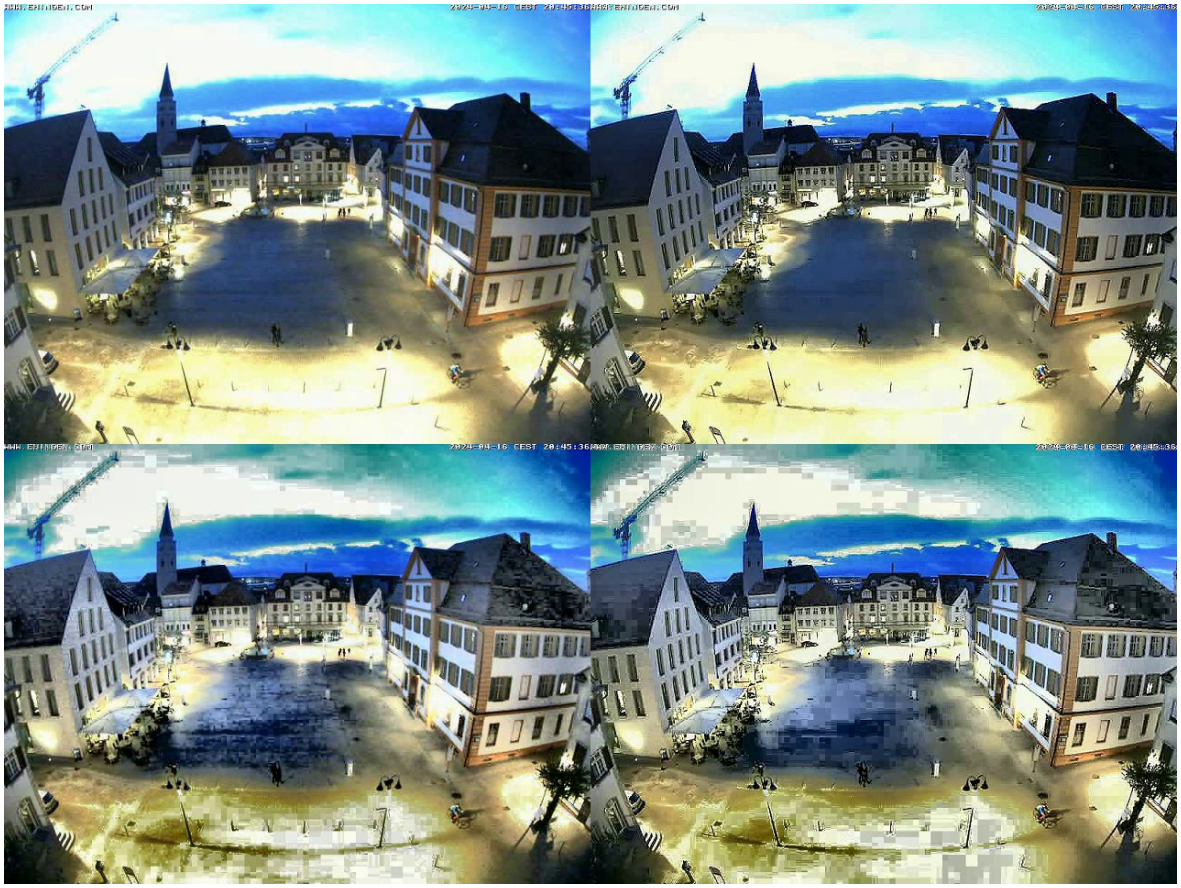


Obr. 3.27: Vľavo hore: zaostrenie odčítaním rozmazaného obrazu, vpravo hore: odstránenie šumu nelokálnym priemerovaním + zaostrenie hranovým filtrom, vľavo dole: zaostrenie odčítaním rozmazaného obrazu + ekvalizácia histogramu, vpravo dole: odstránenie šumu nelokálnym priemerovaním + zaostrenie hranovým filtrom + ekvalizácia histogramu

nie obrazu šumu nelokálnym priemerovaním. Vtedy je vhodnejšie zvoliť zaostrovanie odčítaním rozmazaného obrazu, než hranovým filtrom. Pre porovnanie uvádzame aj kombináciu tých istých zaostrovacích a odšumovacieho algoritmu s adaptívnou ekvalizáciou v ukážke 3.28.

3.5 Poveternostné vplyvy

Kamery umiestnené v exteriéri budú vo väčšine lokalít občas čeliť podmienkam ako je hmla či dážď. Na video strba_snow_fog sme nasadili algoritmy v snahe nielen zlepšiť detail v obraze, pri ktorom sme sa najviac sústredili na konštrukciu lanoviek v pozadí, keďže ide o asi najjednoduchšie a zároveň najsmerodajnejšie posúditeľný detail, ale aj na potlačenie efektu hmly v scéne. Vyskúšali sme množstvo rôznych nastavení adaptívnej ekvalizácie histogramu, a práve poznatky z tohto videa nás viedli k tomu, že vo vyššie zlepšovaných videách sme volili ako primárnu možnosť okolie 4 pixelov a limit



Obr. 3.28: Vľavo hore: zaostrenie odčítaním rozmazaného obrazu, vpravo hore: odstránenie šumu nelokálnym priemerovaním + zaostrenie hranovým filtrom, vľavo dole: zaostrenie odčítaním rozmazaného obrazu + adaptívna ekvalizácia histogramu, vpravo dole: odstránenie šumu nelokálnym priemerovaním + zaostrenie hranovým filtrom + adaptívna ekvalizácia histogramu

Obr. 3.29: Vľavo hore: zaostrenie hranovým filtrom, vpravo hore: zosvetlenie o hodnotu 50 (v každej RGB zložke), vľavo dole: zaostrenie hranovým filtrom + zosvetlenie, vpravo dole: zosvetlenie + zaostrenie hranovým filtrom

rozdielneho kontrastu 20, a ako sekundárnu, agresívnejšiu alternatívu, okolie 8 pixelov a limit kontrastu 40. Problémom inak zvolených kombinácii a pomerov medzi okolím a limitom jasu boli ešte výraznejšie artefakty v podobe “mozajkovania” obrazu. Avšak všetky ekvalizácie histogramu tu zapríčinili skôr degradáciu obrazu. Odskúšali sme aj potenciálny prínos jednoduchého zosvetlenia obrazu, avšak podľa očakávaní, keďže aplikáciou zosvetlenia pripočítaním konštanty k intenzite každého pixelu nedochádza k žiadnemu zvýrazneniu rozdielov medzi jednotlivými intenzitami, práveže na svetlých miestach dôjde k jeho zániku. Efekt síce na pohľad pôsobí "čistejšie", neprináša však v prípade bezpečnostných kamier žiaden osoh. Ukážku tohto možno vidieť na obrázku 3.29.

Nakoniec sme došli k tomu, že najprínosnejšie je v tomto prípade použitie len samotného zaostrovania hranovým filtrom. Na ukážke 3.30 toto kombinujeme s ekvalizáciami histogramu pred alebo po, ale snímku vpravo hore považujeme za lepšiu než tie, ktoré podstúpili ekvalizácie.

Klip parking bol natočený za dažďa, ten však na kamere prakticky nebadať, a nebude ho badať ani na väčšine bezpečnostných kamier, pokiaľ nepôjde o skutočne hustý dážď, tak ako máme problém zachytiť dážď na kamere smartfónu. Väčšina bezpečnostných kamier má zároveň predĺženú striešku nad objektívom, ktorá ho chráni pred dopadom kvapiek. Čo sa teda vôbec mení vplyvom dažďa na záberoch? Nie veľa, no zároveň jedna vec je kľúčovou a neprekonateľnou prekážkou - dáždniky. Drvivá väč-

Obr. 3.30: Vľavo hore: pôvodný záber, vpravo hore: zaostrenie hranovým filtrom, vľavo dole: zaostrenie hranovým filtrom + ekvalizácia histogramu, vpravo dole: + zaostrenie hranovým filtrom + adaptívna ekvalizácia histogramu

šina exteriérových bezpečnostných kamier je umiestnená vo výške, pri pohľade z ktorej kupola dáždnika kompletne prekrýva tvár človeka v akejkoľvek vzdialenosti, v ktorej by ešte bola rozoznateľná. Toto je prekážka, s ktorou sa nedokážeme vysporiadať. Druhá vec, čo sa mení oproti tomu, keď je obloha len zatiahnutá oblakmi, je voda na zemi. Niektoré povrchy môžu vytvárať značné zrkadlenie, keď sú pokryté hoci aj veľmi plytkou vrstvou vody. V našej doméne to nemá priamočiaru podstatnosť - prakticky všetky objekty, ktorých prítomnosť v sledovanom priestore nás zaujíma, sa pohybujú nad povrchom - avšak keď čierny asfalt prestane byť na zázname čiernym, objekty na ňom nemusia oproti nemu tak vyniknúť, čo by mohlo mierne komplikovať prácu niektorým algoritmom a degradovať kvalitu ich výsledku, napríklad aj algoritmom slúžiacim na automatickú identifikáciu či klasifikáciu objektov, pre ktoré sa v tejto práci tiež snažíme maximalizovať objem a čistotu výstupnej informácie. Došli sme k záveru, že na video dostatočne vysokého rozlíšenia za solídnych svetelných podmienok nemá význam aplikovať ostrenie hranovým filtrom, keďže dôjde k zbytočnému preostreniu, a ostrenie odčítaním rozmazaného obrazu v zmysluplnej miere zas neprodukuje badateľný rozdiel. Ekvalizácie histogramu sú zas zbytočné, ak ide o takýto dobre no hlavne uniformne osvetlený priestor, a práveže pôsobia kontraproduktívne, a to jednak zdôraznením lesku vlhkých plôch, a taktiež zhoršením čitateľnosti farby svetlých a tmavých automobilov, ktorá môže byť v niektorých prípadoch predmetom záujmu. Na ukážke 3.31 vidíme, že veľmi malé zlepšenie oproti originálnemu obrazu môžeme so zaostrením

Obr. 3.31: Vľavo hore: pôvodný záber, vpravo hore: zaostrenie odčítaním rozmazaného obrazu, vľavo dole: zaostrenie odčítaním rozmazaného obrazu + ekvalizácia histogramu, vpravo dole: + zaostrenie odčítaním rozmazaného obrazu + adaptívna ekvalizácia histogramu

odčítaním rozmazaného obrazu dosiahnuť, no originálny záznam je dostatočne kvalitný a nevieme ho vhodnejšie predspracovať, alebo zlepšiť priamo z hľadiska ľudského pozorovateľa. Za povšimnutie stojí kráčajúca osoba vľavo hore s bledými nohavicami, ktorá na ekvalizovaných záberoch splýva s lesknúcim sa vlhkým asfaltom.

3.6 Prehľad implementovaných metód

Nakoniec ponúkame prehľad všetkých testovaných metód, ich časovej náročnosti, závislosti od parametrov nastavených pred alebo pri volaní samotnej metódy, a zhodnotenia ich vhodnosti, v tabuľke 3.2 a 3.3.

kategória	algoritmus	očakávateľné μs/px	závislosť od parametrov	uplatnenie
zvyšovanie rozlíšenia	bikubická interpolácia	<0,1	nie	najrýchlejší spôsob zvýšenia rozlíšenia, nezaostáva za neur. sieťami do neakceptovateľnej miery
	EDSR	235	nie	najkonštruktívnejšia a najschopnejšia
	ESPCN	0,8	nie	najrýchlejšia neurónová sieť, najvhodnejšia, keďže vo veľkej väčšine prípadov nezaostáva ani za EDSR
	FSRCNN	1,8	nie	zmysluplne rýchla alternatíva ESPCN
	LapSRN	8,2	nie	časovo najkonzistentnejšia, vnáša najmenej artefaktov
zaostrenie	hranovým filtrom	<0,1	áno	najefeknejšie pre zvýraznenie hrán, vhodné nasadiť po odstránení šumu nelok. priemerovaním, v kombinácii s neur. sieťami až po ich nasadení
	odčítaním rozmazaného obrazu	<0,1	áno	hladší spôsob zaostrenia, efekt nie je význačný, no netvorí artefakty, nepotrebuje kombinovať s odstraňovaním šumu

Tabuľka 3.2: Tabuľka implementovaných metód - 1. časť

kategória	algoritmus	očakávateľné μs/px	závislosť od parametrov	uplatnenie
ekvalizácia histogramu	globálna v YCrCb formáte	<0,1	nie	vhodná na kompenzáciu vysokého dyn. rozsahu scény, nadbytočná v dobrých svetel. podmienkach, nežiadúca, ak je osoba presvetlená voči tmavému pozadiu
	adaptívna v YCrCb formáte	<0,1	áno	vnáša značné artefakty - efekt mozajkovania, efektívna pre vyzdvihnutie tmavých miest
zvýšenie jasu	pripočítaním konštantnej hodnoty	<0,1	áno	zvyšuje pocitovú čitateľnosť záberu ako takého, no je kontraproduktívne v zmysle zvýraznenia vizuálnej informácie
odstránenie šumu	nelokálnym priemerovaním	34	áno	efektné, vhodné kombinovať so zaostrením hran. filtrom, zachováva detaily do uspokojivej miery, výpočtovo náročné
odstránenie rozmazania	dekonvolúciou vo frekvenčnom spektre + nutná ekvalizácia histogramu	0,22	áno	v našej doméne nasadenia zbytočné a kontraproduktívne
rekolorizácia	neurónovou sieťou Deoldify	12	áno	spoľahlivá, ak záber obsahuje referenčné objekty, inak nevyspytateľná
detekcia pohybu	pomocou IR spektra	<0,01	nie	nápomocné v tmavých priestoroch, kvalita závisí od kvality IR obrazu z kamery

Tabuľka 3.3: Tabuľka implementovaných metód - 2. časť

Záver

V tejto práci sme preskúmali, implementovali a porovnali možnosti zlepšenia kvality a čitateľnosti vizuálnej informácie vo videách z bezpečnostných kamier. Výskum v tejto práci by mal byť priamo nápomocný k integrácii softvérov, ktoré:

- Zvýšia celkovú mieru čitateľnosti priestoru sledovaného ľudským operátorom prostredníctvom systému bezpečnostných kamier v reálnom čase, umožňujúc tak efektívnejšie alebo včasnejšie predchádzať škodnej či kriminálnej činnosti, či riešiť iné problémy objaviace sa v medziach monitorovanej zóny.

- Zvýraznia doménovo relevantnú vizuálnu informáciu pri retrospektívnom prehliadaní a vyšetrovaní záznamu a na ňom (potenciálne) zachyteného kontextu. Aj ak na základe snímok stále nedokážeme priradiť tvár konkrétnemu človeku (napríklad v policajnej databáze), alebo jednoznačne určiť evidenčné číslo vozidla, pretože unikátnu informáciu nezachytenú vo videu nedokážeme dotvoriť bez rovnakej referencie, zlepšenie viditeľnosti čŕt môže potvrdiť či vyvrátiť prítomnosť nejakej osoby či vozidla na mieste, čo má samo o sebe vysokú výpovednú hodnotu pre vyšetrovanie spáchanej aktivity.

- Zdôraznia podkladové informácie pre algoritmy rozpoznávajúce objekty v obraze, a zvyšujú tak pravdepodobnosť správnej klasifikácie, čo zvýši spoľahlivosť automatizovaných procesov v dohľadových systémoch.

Program, ktorý sme pre účely práce napísali, umožňuje priamu aplikáciu všetkých preskúmaných metód a úpravu ich pracovných parametrov na videách vyhotovených bezpečnostnými kamerami v reálnom ostrom nasadení, alebo na živých prenosoch z týchto kamier po pripojení sa na IP adresu kamery prostredníctvom jej pridania do zdrojov priloženej triedy LiveVideo v našom doplnkovom module live_video_acquirement. V tejto písomnej časti práce sme zároveň popísali, ktoré metódy sa hodia v akých scenároch nasadenia.

Obzvlášť prínosným vie byť zaostrovanie hranovým filtrom, ktoré vieme aplikovať na priamy prenos z kamery, na ktorú sme prostredníctvom nášho programu pripojení. Ak spracúvame už nahraný videozáznam, stojí za to video pred zaostrovaním zbaviť šumu bez straty unikátneho detailu metódou nelokálneho priemerovania. Medzi týmto

zbavením šumu a zaostrovaním má tiež zmysel nasadiť jednu z implementovaných neurónových sietí, ak má význam pre nás zväčšenie rozlíšenia obrazu aspoň štvornásobne, kedy začínajú mať neurónové siete kvalitatívny náskok pred bikubickou interpoláciou. Výsledok pri použití najrýchlejšieho modelu ESPCN bude takmer na nerozoznanie od výsledkov modelov FSRCNN a LapSRN. Ak nie sme limitovaní časom, ktorý môže pri nasadení tohto modelu už na niekoľkosekundové video vo vysokom rozlíšení dosahovať hodiny, o trochu lepší výsledok vieme niekedy dosiahnuť použitím modelu EDSR. Ak sa vrátime k zaostrovaniu, a chceme ho robiť v reálnom čase, naživo, a neprajeme si preostrenie vnášajúce do obrazu artefakty, zaostrovanie odčítaním rozmazaného obrazu vie priniesť pozitívne výsledky bez nutnosti predspracovania. Ekvalizácie histogramov nám pomôžu zviditeľniť dianie v presvetlených a tmavých regiónoch pri záberoch s vysokým dynamickým rozsahom, pričom silu efektu adaptívnej ekvalizácie vieme regulovať nastavením jej parametrov. Čím agresívnejšie sa pokúsime vytiahnuť malé rozdiely v zachytených intenzitách, tým viac artefaktov do obrazu vnesieme, značne znižujúcich prehľadnosť, a komplikujúcich prácu pre prípadné nasadenie neurónových sietí alebo rozpoznávacích algoritmov na výsledok adaptívnej ekvalizácie. Preto sú ekvalizácie histogramu situačné. Rovnako situačná je aj implementovaná rekolorizácia, ktorá je spoľahlivá len ak sú na šedotónovom zábere (napríklad z kamery kombinujúcej obraz zo spektier viditeľného svetla a infračerveného žiarenia) obsiahnuté objekty ktorým obvykle prináležia nejaké typické farebné oditene - napríklad zeleň, potenciálne obloha kontrastujúca s oblakmi. Tam kde je rekolorizácia zmysluplne použiteľná, umožňuje farebne separovať objekty v obraze od pozadia a zlepšiť tak ich identifikovateľnosť oproti šedotónovému obrazu. Aplikovať ju však dokážeme len na už nahrané video, kvôli časovej náročnosti. Taktiež náš program s jednoduchou úvodnou asistenciou od ľudského pozorovateľa umožnuje pre kamery s paralelným záznamom v infračervenom spektre pomocou neho vyznačovať približné siluety pohybujúcich sa osôb v RGB zázname v ľubovoľne nepriaznivých svetelných podmienkach. Vstupom do tejto metódy môže byť aj živý prenos v týchto dvoch spektrách.

Ďalší vývoj v tejto oblasti sľubuje príchod ďalších podstatných zlepšení. Rozpracovaných, alebo v stave pred zavedením do produkcie, je momentálne viacero algoritmov, vnášajúcich nové možnosti do našej oblasti. Sú to napríklad neurónové siete špecializované na zvyšovanie rozlíšenia videa využívajúc kontext medzi jednotlivými snímkami, softvérové riešenia kombinujúce denné a nočné zábery tej istej scény pre spoľahlivejšiu rekolorizáciu za nízkej viditeľnosti, a množstvo navrhovaných zlepšení už existujúcich, aj tu použitých algoritmov, ako napríklad modifikácie ekvalizácie histogramu, ktoré takpovediac len čakajú, kým ich autori zverejnia v nasaditeľnej forme. V najbližších rokoch budú pravdepodobne čoraz väčšiu rolu aj v doméne bezpečnostných kamier hrať predtrénované hlboké neurónové siete, i keď kompromis medzi kvalitou a rýchlosťou umožňujúcou nasadenie na prenos obrazov v reálnom čase sa zdá byť zatiaľ veľkou

výzvou. Tak či tak sa budúcnosť identifikácie objektov v priestoroch pod dohľadom bezpečnostných kamier javí vďaka veľkému záujmu výskumníkov o oblasť spracovania obrazu nádejne.

Literatúra

[1] R. Huber, "Slowenien, portorose, salzlagerhalle am hafen (hdr-aufnahme vs. normalbelichtung)," 2008. [Online; accessed 22-March-2023].

[2] F. Brandi, R. de Queiroz, and D. Mukherjee, "Super resolution of video using key frames," in *2008 IEEE International Symposium on Circuits and Systems*, pp. 1608–1611, 2008.

[3] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Deblurring images via dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 10, pp. 2315–2328, 2018.

[4] T. Soumya and S. Thampi, "Self-organized night video enhancement for surveillance systems," *Signal, Image and Video Processing*, vol. 11, 01 2017.

[5] A. Buades, B. Coll, and J.-M. Morel, "Non-Local Means Denoising," *Image Processing On Line*, vol. 1, pp. 208–212, 2011. `https://doi.org/10.5201/ipol.2011.bcm_nlm`.

[6] A. Sakaushi, K. Kanai, J. Katto, and T. Tsuda, "Edge-centric video surveillance system based on event-driven rate adaptation for 24-hour monitoring," in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 651–656, 2018.

[7] Y. Rao and L. Chen, "A survey of video enhancement techniques," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 3, pp. 71–99, 01 2012.

[8] Y. Ali, W. Md-Esa, A. Malik, and M. N. MOHAMAD SAAD, "Tone mapping of hdr images: A review," pp. 368–373, 06 2012.

[9] M. Nilsson, "Ultra high definition video formats and standardisation," 04 2015.

[10] D. Kong, M. Han, W. Xu, H. Tao, and Y. Gong, "Video super-resolution with scene-specific priors.," in *BMVC*, pp. 549–558, 2006.

[11] Y. Liu, A. Wong, and P. Fieguth, "A structure-guided conditional sampling model for video resolution enhancement," in *2011 18th IEEE International Conference on Image Processing*, pp. 1169–1172, 2011.

[12] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[13] W. Freeman, T. Jones, and E. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56–65, 2002.

[14] F. Cristina, S. H. Dapoto, C. C. Russo, and O. N. Bria, "High resolution images from low resolution video sequences," *Journal of Computer Science & Technology*, vol. 5, 2005.

[15] R. Schultz and R. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 996–1011, 1996.

[16] X. Ge, J. Tan, and L. Zhang, "Blind image deblurring using a non-linear channel prior based on dark and bright channels," *IEEE Transactions on Image Processing*, vol. 30, pp. 6970–6984, 2021.

[17] S. Chunhe, Z. Hai, and J. Wei, "Motion deblurring from a single image using multi-layer statistics priors," in *2011 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 481–482, 2011.

[18] X. Wang, K. C. Chan, K. Yu, C. Dong, and C. Change Loy, "Edvr: Video restoration with enhanced deformable convolutional networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019.

[19] X. Ji, Y. Cao, Y. Tai, C. Wang, J. Li, and F. Huang, "Real-world super-resolution via kernel estimation and noise injection," in *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2020.

[20] Y. Tian, Y. Zhang, Y. Fu, and C. Xu, "Tdan: Temporally-deformable alignment network for video super-resolution," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[21] J. Pan, S. Cheng, J. Zhang, and J. Tang, "Deep blind video super-resolution," 2020.

[22] D. Su, H. Wang, L. Jin, X. Sun, and X. Peng, "Local-global fusion network for video super-resolution," *IEEE Access*, vol. 8, pp. 172443–172456, 2020.

[23] Y. Jo, S. W. Oh, J. Kang, and S. J. Kim, "Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[24] MMagic Contributors, "MMagic: OpenMMLab multimodal advanced, generative, and intelligent creation toolbox." `https://github.com/open-mmlab/mmagic`, 2023.

[25] R. Taiyenjam, O. Imocha Singh, K. Singh, T. Sinam, and T. Singh, "Image enhancement by adaptive power-law transformations," 12 2010.

[26] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," in *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pp. 661–670, 2023.

[27] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," in *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pp. 671–678, 2023.

[28] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," in *Computer graphics forum*, vol. 22, pp. 419–426, Wiley Online Library, 2003.

[29] R. Mantiuk and H.-P. Seidel, "Modeling a generic tone-mapping operator," in *Computer Graphics Forum*, vol. 27, pp. 699–708, Wiley Online Library, 2008.

[30] B. H.K and K. N, "Video enhancement using histogram equalization with jnd model," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, pp. 2506–2511, 07 2019.

[31] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," 2017.

[32] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," 2016.

[33] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," 2016.

[34] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Fast and accurate image super-resolution with deep laplacian pyramid networks," 2018.

[35] I. Žeger, S. Grgic, J. Vuković, and G. Šišul, "Grayscale image colorization methods: Overview and evaluation," *IEEE Access*, vol. 9, pp. 113326–113346, 2021.

[36] T. Soumya and S. M. Thampi, "Recolorizing dark regions to enhance night surveillance video," *Multimedia Tools and Applications*, vol. 76, pp. 24477–24493, 2017.