VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

REIDENTIFIKACE AUTOMOBILŮ VE VIDEU

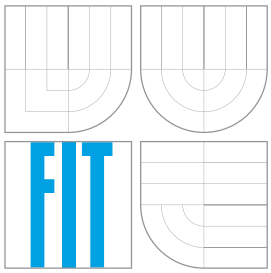
BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

DOMINIK ZAPLETAL

BRNO 2015

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

REIDENTIFIKACE AUTOMOBILŮ VE VIDEU

VEHICLE RE-IDENTIFICATION IN VIDEO

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE — DOMINIK ZAPLETAL
AUTHOR

VEDOUCÍ PRÁCE — doc. Ing. ADAM HEROUT, Ph.D.
SUPERVISOR

BRNO 2015

Abstrakt

Tato bakalářská práce se zabývá problémem reidentifikace automobilů ve videu. Reidentifikace automobilů ve videu je založena na porovnávání částí obrazu získaného z různých kamer. Tato práce je zaměřena zejména na reidentifikaci automobilů samotnou a vychází z předpokladu, že problém detekce automobilů ve videu je již vyřešen v podobě vytvořeného 3D ohraničujícího kvádru kolem vozidla. Problém reidentifikace je vyřešen pomocí barevných histogramů, histogramů orientovaných gradientů a lineárního regresoru. Příznaky jsou používány v oddělených modelech za účelem dosažení nejlepších výsledků v nejkratším výpočetním čase procesoru. Navrhovaná metoda pracuje s vysokou přesností (60 % opravdu pozitivních rozpoznání s 10 % mírou falešně pozitivních případů na náročné datové sadě) s výpočetním časem procesoru (Core i7) 85 milisekund pro jednu reidentifikaci vozidla za předpokladu video vstupu v plném HD rozlišení. Použitím této práce v distribuovaných dopravních monitorovacích systémech je možné zjistit důležité parametry jako doba cestování, směry dopravních toků nebo dopravní informace.

Abstract

This thesis deals with the vehicle re-identification in video problem. Vehicle re-identification is based on matching image parts obtained from different cameras. This work is focues on the re-identification itself assuming that the vehicle detection problem is already solved including extraction of a full-fledged 3D bounding box. The re-identification problem is solved by using color histograms, histograms of oriented gradients by a linear regressor. The features are used in separate models in order to get the best results in the shortest CPU computation time. The proposed method works with a high accuracy (60 % true positives retrieved with 10 % false positive rate on a challenging subset of the test data) in 85 milliseconds of the CPU (Core i7) computation time per one vehicle re-identification assuming the Full HD resolution video input. The applications of this work include finding important parameters like travel time, traffic flow, or traffic information in a distributed traffic surveillance and monitoring system.

Klíčová slova

reidentifikace automobilů, anonymní reidentifikace, počítačové vidění, regrese, extrakce příznaků, porovnávání obrazu, strojové učení

Keywords

vehicle re-identification, anonymous re-identification, computer vision, regression, feature extraction, image matching, machine learning

Citace

Dominik Zapletal: Reidentifikace automobilů ve videu, bakalářská práce, Brno, FIT VUT v Brně, 2015

Reidentifikace automobilů ve videu

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana doc. Ing. Adama Herouta, Ph.D. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Dominik Zapletal
19. května 2015

Poděkování

Děkuji vedoucímu práce panu doc. Ing. Adamovi Heroutovi, Ph.D. za přátelský přístup na konzultacích a cenné rady, návrhy a připomínky, které napomohly úspěšnému dokončení této práce a článků na studentskou konferenci Excel@FIT 2015 a konferenci ITSC 2015. Dále bych chtěl poděkovat paní Ing. Markétě Dubské, Ph.D. za ochotu zpracovat detekčním systémem natočená videa.

Obsah

Kapitola 1

Úvod

Aktuální dopravní informace jsou v dnešní době čím dál více požadovány a to z mnoha důvodů: pro sbírání statistických dat [22], pro okamžité ovládání dopravních signálů [19], pro prosazování práva [16, 27], pro redukci dopravy a optimalizaci dopravních toků [6], pro výpočet trasy v GPS navigacích [15], pro detekci nebezpečných událostí na silnicích včetně dopravních nehod, stojících vozidel a překážek [1], atd. Jeden ze způsobů, jak takové informace získat, je reidentifikace automobilů. V mnoha dopravně frekventovaných oblastech jsou již zabudovány různé druhy pouličních kamer, proto by bylo výhodné využít zdroje, které jsou již k dispozici, bez nutnosti jejich výměny nebo modifikace. Pro dosažení maximální univerzálnosti systému je žádoucí, aby byla reidentifikace založena pouze na vizuálních vlastnostech vozidla [24].

Tato práce se zabývá problematikou reidentifikace automobilů ve videu. Pod tímto pojmem se rozumí extrakce dostatečného množství informací z různých video zdrojů (záznam, webkamera, ...) a efektivní využití těchto informací k nalezení dvojic totožných automobilů v různých množinách. Řešení problému reidentifikace v této práci popsané je založeno pouze na vizuálních charakteristikách vozidla. Jedná se tedy o anonymní reidentifikaci, jejímž cílem je nalezení daného vozidla pouze na základě jeho geometrických a barevných vlastností. Anonymní reidentifikace nabízí univerzální využití a nezávislost na registrační značce automobilu, která je v obraze mnohdy těžko dosažitelná. Díky tomu nemusí být kladeny otázky týkající se ochrany soukromí.

Práce je rozdělena do několika kapitol. V kapitole 2 jsou rozebírány některá z řešení zabývající se problematikou reidentifikace v různých oblastech včetně zhodnocení jejich silných a slabých stránek. Dále je v kapitole 3 uveden teoretický základ některých metod, které v počítačovém vidění slouží pro výpočet příznaků a strojové učení. Tyto metody jsou zároveň relevantní k řešení daného problému. Návrh systému, který by měl reidentifikaci automobilů ve videu řešit, je uveden spolu s konkrétní implementací v kapitole 4. Způsob vyhodnocení a dosažené výsledky z implementovaného systému jsou nakonec zobrazeny a popsány v kapitole 5.

Řešení problému reidentifikace automobilů ve videu v této práci popsané bylo prezentováno na studentské konferenci inovací, technologií a vědy v IT Excel@FIT 2015, kde bylo ohodnoceno cenou společnosti FEI. Zároveň byl poslán článek na konferenci IEEE - ITSC (Intelligent Transportation Systems Conference) 2015, který je v době dokončení této práce stále v recenzním řízení.

Kapitola 2

Existující práce zabývající se reidentifikací

Problematikou reidentifikace se v posledních dvou desetiletích zabývá nemalé množství prací. Reidentifikace se netýká bezvýhradně jen automobilů, ale také lidí a jiných objektů. Zároveň existují různé přístupy k řešení tohoto problému, přičemž každý z nich má svoje klady a zápory. I když se řešení liší podle konkrétních požadavků na systém, často vycházejí ze stejných principů. V této kapitole jsou uvedeny některé přístupy, které problém reidentifikace nebo jeho část řeší, včetně zhodnocení jejich silných a slabých stránek.

2.1 Detekce automobilů v obraze

Jedním z důležitých kroků pro reidentifikaci automobilů je právě jejich detekce v obraze. Čím přesnější detekci automobilu je schopen systém provést, tím vyšší je pravděpodobnost úspěchu samotné reidentifikace, jelikož výběr správné oblasti v obraze, kde se automobil nachází, je pro přesnost reidentifikace směrodatný. I když se tato práce v řešení problému reidentifikace detekcí automobilů v obraze nezabývá, je vhodné si některé základní principy k řešení tohoto problému uvést.

Postup detekce automobilů v obraze ze stacionární kamery se dělí na dva hlavní kroky. V prvním kroku jsou používány metody pro výběr oblasti zájmu (ROI). V druhém kroku pak probíhá samotná detekce automobilu pomocí různých metod [29].

Metod pro výběr oblasti zájmu je několik a dělí se většinou do tří kategorií: metody odčítání snímků, metody aktualizace pozadí a metody virtuální smyčky.

Metody odčítání snímků jsou založeny na rozdílech pixelů nebo bloků v obraze mezisnímků (inter-frames). Tyto techniky využívají rozdílů dvou nebo více po sobě jdoucích snímků za účelem nalezení oblasti zájmu. Výhodou těchto metod je jejich přizpůsobivost. Nevýhodou je velká závislost na časovém intervalu po sobě následujících snímků a také na rychlosti pohybu vozidla v obraze.

Metody aktualizace pozadí jsou často používané pro detekci pohybu v mnohých aplikacích pracujících v reálném čase pro detekci a sledování vozidel v obraze. Patří mezi ně například metoda průměrného snímku, metoda selektivního aktualizování, metoda minimální a maximální hodnoty intenzity a kombinace Gaussovy metody a techniky k-means shlukování. Princip těchto metod spočívá ve vypočítání rozdílu následujícího snímku se snímkem zkonstruovaného pozadí. Tímto způsobem jsou detekovány objekty v popředí obrazu. Po následném vyjmutí pozadí jsou získány kompletní informace o těchto objektech. Použití

zmíněných technik je problematické ve scénách, ve kterých se dynamicky mění osvětlení.

Metody virtuální smyčky využívají koncept indukční smyčky pro detekování projíždějícího vozidla monitorováním změn osvětlení v přednastavené oblasti snímku. Tato oblast je vybrána člověkem a odvíjí se od požadavků na detekci v konkrétní scéně. Jelikož jsou zpracovávány jen určité oblasti snímků, je výpočetní náročnost těchto metod nízká. Jejich parametry je ovšem složité správné nastavit. Funkce těchto technik je omezená.

Vstupní informace pro metody detekce automobilů v obraze je množina oblastí zájmu získaná některou z doposud zmíněných metod. Detekční metody se také podle Wanga *aj.* [29] rozdělují do tří kategorií: metody založené na znalostech, metody založené na pohybu a tzv. vlnkové metody (wavelet-based methods).

Metody založené na znalostech vycházejí z předchozích znalostí a rozhodují se, zda-li je vstupní ROI vozidlo nebo ne.

Metody založené na pohybu vycházejí z výpočtu optického toku. Body ve snímcích se zdají být v pohybu v důsledku relativního posunu snímače a snímané scény. Vektorové pole tohoto pohybu je označováno jako optický tok. Tyto metody používají charakteristiky toků vektorů pohybujících se objektů v čase pro detekci pohybujících se oblastí v sekvenci snímků. Pomocí optických toků je možné detekovat nezávisle se pohybující vozidla v obraze. Nevýhodou je jejich vysoká citlivost na šum. Aby bylo možné těmito metodami zpracovávat obraz v reálném čase, je většinou nutné použít specializovaného hardware.

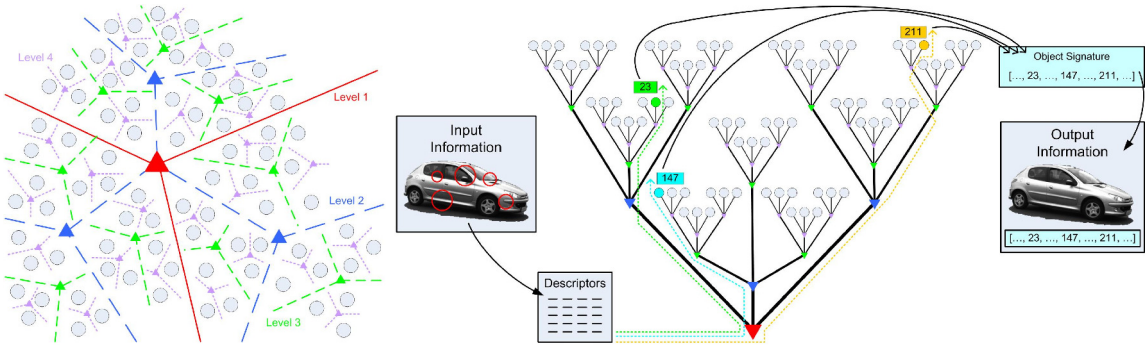
Vlnkové metody dosahují v detekci automobilů velice dobrých výsledků při obdržení dočasných prostorových informacích o pohybu vozidla. Jejich výhoda je nižší výpočetní náročnost a jednoduchá implementace. Nicméně jsou i tyto metody poměrně citlivé na šum v obraze a různorodost časování pohybů.

2.2 Reidentifikace založená na vizuálních charakteristikách objektu

Reidentifikační systémy založené na vizuálních charakteristikách daného objektu (automobil, člověk, letadlo, ...) jsou jedny z nejrozšířenějších. Hlavní výhody takových systémů jsou jejich univerzální použití, anonymita a větší odolnost vůči změnám v obraze v podobě rotace, měřítka, šumu a osvětlení. Proto není nutné provádět strojové učení na konkrétních podmínkách. Díky anonymitě není získána (ve většině případů ani nemůže být) identita reidentifikováného objektu. Nevýhodou je jejich menší přesnost a vyšší paměťová a výpočetní náročnost. V následujících odstavcích budou uvedeny některé ze zajímavých řešení problému reidentifikace vycházející pouze z vizuálních charakteristik objektu v obraze.

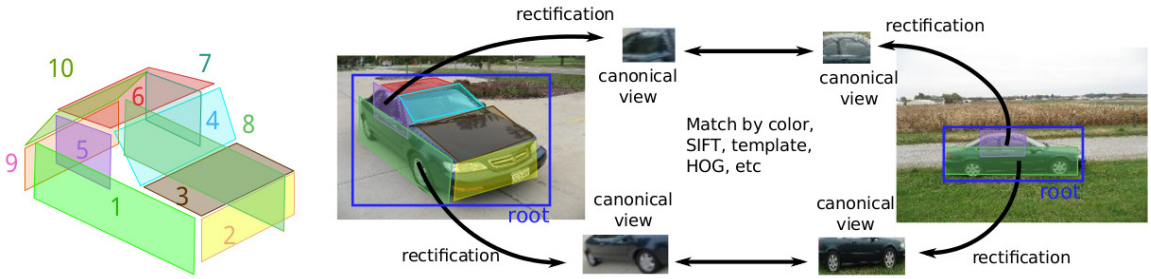
Arth *aj.* [2] vytvořili autonomní decentralizovaný systém pro reidentifikaci automobilů ve videu fungující v reálném čase. Systém reprezentuje síť vzájemně se nepřekrývajících chytrých kamer, které mezi sebou nezávisle komunikují. Kamery není potřeba kalibrovat na konkrétní scénu, což je velice pozitivní vlastnost, jelikož správná kalibrace je poměrně náročný proces zejména z důvodu, že se daná scéna může v čase značně měnit (roční období, denní doba, počasí, ...). Pro extrakci lokálních rysů automobilu používají DOG (difference of Gaussian) detektor, který jim poskytuje přesné klíčové body, které se často opakují. DOG detektor je založen na principu rozdílů Gaussově rozostřených obrazů. Například aplikováním DOG filtru na obraz je dosaženo detekce hran objektů ve fotografii (obrázek 2.1). Klíčové body vypočítané detektorem vypovídají o struktuře vstupního obrazu. Výhoda DOG detektoru je, že je téměř úplně založen jen na zmíněné filtraci obrazu a na operacích sčítání a odčítání, což znamená nízkou náročnost na výpočetní výkon. Při

Obrázek 2.1: Ukázka DOG filtru aplikovaného na obraz. Vlevo je originální obraz a vpravo obraz po aplikaci filtru. Díky DOG filtru vynikly hrany, celková struktura a textury obrazu. DOG filtr byl aplikován programem GIMP.



Obrázek 2.2: Ilustrace ukazující mechanismus generování podpisu detekovaného objektu (automobil) v obraze, jehož výsledkem je tzv. slovníkový strom. Obrázky převzaty z [2].

správně nastavených parametrech vyhledávání automobilů a filtrování obrazu je jejich algoritmus s použitím DOG detektoru také méně paměťově náročný. U vestavěných systémů (například zmíněné kamery) je efektivní využití paměti jeden z hlavních požadavků na algoritmy. Filtraci obrazu implementovali v pevné řádové čárce. Díky tomu podstatně zvýšili rychlost výpočtu. DOG detektor, uvedený v jejich práci, vypočítá několik klíčových bodů v různém měřítku po ploše automobilu v obrazu. Na základě vypočítaného bodu pak vyextrahují kruhovou výseč, jejíž poloměr vychází z měřítka klíčového bodu. Z těchto kruhových výsečí jsou vypočítány hlavní příznaky automobilu. Pro výpočet příznaků používají PCA-SIFT deskriptor [18]. Velký počet vypočítaných příznaků je pak kvantován do opakujících se k-means shluků. Každý ze shluků je dělen v uzlech tak dlouho, dokud je dělení možné. Data tímto způsobem vypočítaná vytváří tzv. slovníkový strom (vocabulary tree), nebo-li podpis (signature). Ilustrace slovníkového stromu a způsob jeho vytvoření je na obrázku 2.2. Množství dat, které je potřeba uložit závisí na dimenzi PCA-SIFT deskriptoru. Slovníkový strom je pak dál optimalizován za účelem snížení jeho velikosti a rychlosti jeho procházení. Velikost příznakových vektorů vypočítaných deskriptorem se pohybuje v rámci několika stovek příznaků. Díky tomu je možno vektory poměrně efektivně porovnávat. Na základě Euklidovy vzdálenosti příznaků v sousedních listech stromu jsou pak odstraňovány nechtěné příznaky náležící zejména k pozadí za automobilem. V jejich systému používají pouze černobílý obraz a ztrácí tak informaci o barevných vlastnostech vozidla, což může způsobit



Obrázek 2.3: Ilustrace principu ko-detekce (reidentifikace) objektů v obraze. Automobil je snímán z různých úhlů a v různém prostředí. Po jeho detekci je rozdělen do částí podle jeho geometrických vlastností a následně jsou příslušné části různými mechanismy (barva, SIFT, HOG, ...) porovnávány mezi sebou. Obrázky převzaty z [4].

zbytečně velké množství falešně pozitivních reidentifikací. Jejich systém byl schopen správně reidentifikovat 87 % automobilů. Nutno dodat, že se jednalo o sestavu sousedních na sebe navazujících kamer. Nebylo tedy vyhledáváno v příliš velké množině automobilů.

V další práci úzce související s problémem reidentifikace se zabývá Bao *aj.* [4] tzv. ko-detekcí objektů v obraze. Jejich systém hledá souvislosti mezi objekty, které jsou pozorovány z různých úhlů. Výsledkem je nalezení stejných instancí objektů včetně zjištění transformace pohledu. Systém vychází z předpokladu, že má pozorovaný objekt neměnný vzhled. Detekovaný objekt v obraze je reprezentován jako 3D geometrické těleso rozdělené do několika částí. Stanovení identity jednotlivých objektů je pak založeno na porovnávání jejich geometrie a vzhledu dílčích částí (obrázek 2.3). Díky tomu je systém schopen nezávisle na pozorovacím úhlu s relativně vysokou přesností určit identitu objektů nalezených v jiném zdroji obrazu. Ko-detekce objektů v obraze zdaleka není triviální problém, jelikož se může vzhled objektu velice měnit v závislosti na úhlu pohledu na daný objekt. Zároveň mohou být určité části viditelné pouze z určitého pozorovacího úhlu. Objekt se také může nacházet v různých prostředích, proto je třeba brát v úvahu jiný druh okolí (pozadí) objektu.

Pouliční kamery jsou použity také k reidentifikaci lidí. Reidentifikace lidí je v některých případech komplexnější problém a to z několika důvodů: při reidentifikaci v rozsáhlejší síti kamer dochází k větším rozdílům osvětlení v různých prostorách, lidská postava za pohybu a z různých úhlu více mění svoji podobu, pozorovaný člověk může změnit svůj vzhled změnou oblečení (svleče si například bundu), atd. Všechny tyto případy nemají triviální řešení.

Reidentifikace lidí má mnohá využití. V obchodním sektoru je často využívána pro sledování a kontrolu zákazníka. Dále se podobné systémy využívají na letištích, v podzemních parkovištích, v nemocnicích pro detekci neobvyklých událostí nebo v různých ústavech pro sledování pacientů. Práce, které reidentifikaci lidí řeší, se liší podle počtu použitých kamer v systému, podle typu kamer (barevné, černobílé, CCD, web kamery, ...), jejich rychlosti a rozlišení, podle typu prostředí, velikosti pozorované oblasti a podle pozice kamer (překrývající se nebo více vzdálené od sebe).

Aziz *aj.* [3] založili svůj reidentifikační systém na klasifikaci vzhledu a segmentaci částí siluety postavy. Detekovaný člověk v obraze je nejdříve klasifikován podle jeho pozice ke kameře. Výsledkem klasifikace je tedy odpověď, zda-li je postava otočena čelem ke kameře nebo naopak. Zjištění této informace napomáhá detektor obličejové části postavy. Aziz *aj.* tvrdí, že díky zmíněné klasifikaci je systém více odolný vůči změnám v osvětlení a obleční. Dalším krokem je pak rozdělení siluety postavy do několika horizontálních částí (hlava, tělo

a nohy), se kterými se následně pracuje. Z dílčích části siluety jsou pak vypočítány příznaky pomocí SIFT a SURF deskriptoru a pomocí tzv. rotačních zájmových bodů v obraze. Jednotlivé části jsou následně srovnávány metodou založenou na EMD (earth mover's distance). Díky porovnávání jen k sobě náležících částí tak snížili množství falešných reidentifikací.

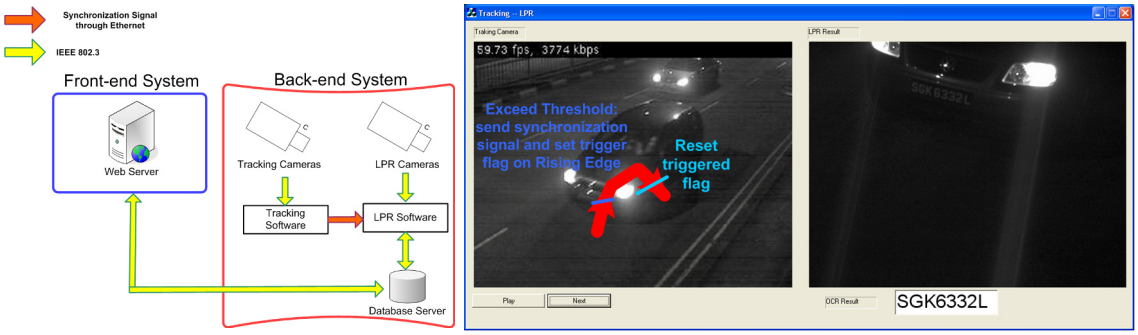
2.3 Reidentifikace automobilů založená na detekci a rozpoznání registrační značky

Další přístupy jsou založeny na detekci a extrakci dat z registrační značky automobilů. Řešení vycházející z této informace dosahují v ideálních podmínkách velice dobrých výsledků, protože je možno registrační značku považovat za unikátní informaci pro daný automobil. Při úspěšné extrakci potřebných dat je pak možno velice jednoduše reidentifikovat vozidla i v rozsáhlých kamerových sítích a databázích s velkým množstvím vozidel, jelikož se v podstatě jedná o porovnávání textových řetězců. Pro dosažení dobrých výsledků jsou ovšem na systém kladeny poměrně velké nároky, což je jedna z hlavních nevýhod, pomineme-li otázku ochrany soukromí. Další negativní vlivy, se kterými se mohou tato řešení potýkat jsou následující:

- Rozlišení zdroje může být nedostačující a proto vstupní data ani informaci o registrační značce neobsahují.
- Povětrnostní podmínky (mlha, silný déšť, sníh, ...) a denní doba (přímé sluneční paprsky směřující do objektivu kamery, záběry v noci/šeru, ...) mohou znesnadňovat čitelnost registračních značek a zabránit tak její detekci a následné extrakci.
- Některé dopravní kamery nemusí být umístěny za účelem dobré viditelnosti registračních značek a jejich čitelnost je tedy nízká.
- Automobil nemusí mít registrační značku čitelnou například z důvodu nečistot (prach, sníh, ...).

Pokud vezmeme v úvahu i jen částečné splnění některé ze zmíněných podmínek, bylo by pro detekční a extrakční systém velice obtížné spolehlivě a rychle rozpoznat jednotlivé registrační značky vozidel. Proto vytvořit univerzální a rychlý reidentifikační systém založený na tomto kritériu (i jen částečně) není triviální.

Wei-Khing *aj.* [13] založili svůj reidentifikační systém na síti dvojic kamer (obrázek 2.4) zabudovaných na dálnicích, který v reálném čase detektuje automobily a rozpoznává jejich registrační značky. Jedna kamera mající menší rozlišení a větší úhel záběru detekuje a sleduje jednotlivá vozidla (TC – tracking camera) a druhá tzv. LPR kamera (licence plate recognition camera), která má vyšší rozlišení a je z pravidla nasměrována záběrem tak, aby co nejlépe zaznamenala pouze registrační značku vozidla. Kamery v jejich systému zaznamenávají jak přední tak zadní část daného automobilu. Zpracovávání videa z kamery s menším rozlišením je méně výpočetně náročné. Avšak rozlišení dostačuje na to, aby byl automobil v obraze detekován a tak byla včas informována LPR kamera. Detekce automobilů v obraze natočené TC kamerou vychází ze dvou modelů. První model se nazývá adaptivní modelování a segmentace pozadí. Tento modul poskytuje detekci automobilu v nekonzistentním venkovním prostředí. Využívají v něm metody PCR (principal color representation), která je schopna přizpůsobivě absorbovat změny v pozadí. Hlavní důvod, proč bylo nutno implementovat modelování pozadí, byla odolnost systému vůči změnám ve scéně (déšť, slunečné



Obrázek 2.4: Systém reidentifikující automobily založený na dvojici kamer TC (tracking camera) a LPR (licence plate recognition camera). Obrázky převzaty z [13].

počasí, oblačno, ...). Pokud taková změna nastane, tak díky adaptivnímu modelování pozadí dojde k aktualizaci pixelů v pozadí a díky tomu se sníží množství falešných detekcí. Pohybující se vozidla v popředí obrazu jsou pomocí segmentace extrahovány a poslány do modulu detekce objektů v popředí obrazu. V tomto modulu používají dva algoritmy. Jeden slouží pro detekci automobilu v obraze ve dne a druhý v noci. Přepínání mezi těmito moduly probíhá automaticky na základě jasových histogramů obrazu. Dvou algoritmů je potřeba z důvodu, že extrakce automobilu v noci předchozím modulem není tak dokonalá jako ve dne. Automobil je detekován, jakmile v obraze dosáhne tzv. spouštěcí zóny (triggering zone). Tato zóna je nakonfigurována specificky pro konkrétní dvojici TC a LPR kamer. Díky detekci automobilů ve stejné oblasti obrazu je možno vyvodit informaci o jeho velikosti. Na základě velikosti vozidla je pak aplikován příslušný filtr redukce šumu. Dále je spouštěcí zóna použita k včasnému informování LPR kamery pro rozeznání registrační značky. Při detekci dochází ke dvěma událostem. Wei-Khing *aj.* je nazývají jako detekce při nástupné hraně a detekce při sestupné hraně. Jakmile automobil vstupuje do spouštěcí zóny dochází k detekci na nástupné hraně, kdy je synchronizována dvojice předních TC a LPR kamer. LPR kamera zaznamená registrační značku automobilu v okamžiku, kdy se vyskytne v zóně pro vyfotografování. Obdobný postup je aplikován u druhé události, která se ovšem vztahuje na dvojici zadních TC a LPR kamer. Implementace jejich reidentifikačního systému automobilů založeného na rozpoznávání registračních značek dosáhla 98 % přesnosti, což je u úlohy tohoto typu velice dobrý výsledek. Je ovšem nutné brát v potaz fakt, že uvedený systém je plně závislý na speciálních kamerách, jejichž instalace je v širším měřítku velice finančně náročná.

Existuje mnoho dalších prací zabývající se reidentifikací automobilů na základě rozpoznávání registrační značky [1, 20, 21]. Ve většině případů vycházejí ze stejného principu: detekce vozidla kamerou s nižším rozlišením, vyfocení oblasti s registrační značkou kvalitnější kamerou, lokalizace registrační značky v obraze a následné rozpoznání jednotlivých písmen a číslic. Tyto systémy jsou většinou aplikovány v případech, kdy je kladen velký důraz na přesnost a správnost výsledků (například mýtné brány nebo úsekové sledování rychlosti automobilů na dálnicích).

Kapitola 3

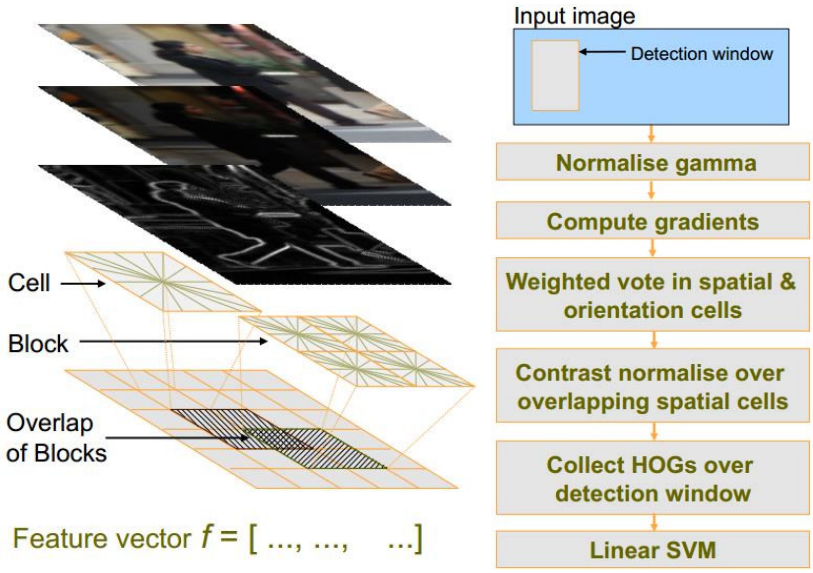
Vybrané metody pro strojové učení a extrakci příznaků

V této kapitole jsou popsány některé z metod strojového učení a extrakce příznaků detekovaných objektů v obraze. V oblasti počítačového vidění existuje mnoho metod v této kapitole neuvedených. Byly vybrány pouze metody relevantní k řešení problému reidentifikace automobilů uvedeném v kapitole 4.

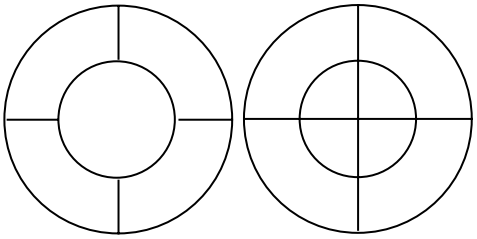
3.1 Histogram orientovaných gradientů

Histogram orientovaných gradientů (HOG) poprvé popsal Navneet Dalal a Bill Triggs ve své práci *Histogramy orientovaných gradientů pro detekci lidí* v roce 2005 [8]. Základní myšlenkou histogramů orientovaných gradientů je, že lokální vzhled objektu může být často dobře charakterizován intenzitami gradientů a směry hran. To je docíleno rozdělením obrazu na regiony (buňky) a následným výpočtem 1D histogramu směrů gradientů a orientací hran přes všechny obrazové body daného regionu. Pro dosažení vyšší přesnosti mohou být místní histogramy normalizované podle kontrastu. Normalizace je aplikována v rámci větší oblasti obrazu (bloky), kde je vypočítána míra intenzity jasové složky. Z vypočítané míry jsou pak normalizovány všechny histogramy vypočítané z buněk spadajících do daného bloku. Díky normalizaci je dosaženo lepších výsledků u obrazových vstupů obsahující větší změny v osvětlení nebo stínech. Na rozdíl od jiných metod pro extrakci příznaků, HOG nevyžaduje předzpracování obrazu (například pro normalizaci barev nebo gama hodnot). Tím je dosaženo ušetření výpočetního výkonu. Uvedený postup je ilustrován na obrázku 3.1.

Jak Dalal *aj.* [8] uvádí ve své práci, pro výpočet gradientů v rámci dané buňky je vhodné použít 1D diskrétní derivační masku $[-1, 0, 1]$. Při použití větších masek dochází k snižování výkonu. Následně je vypočítána prostorová orientace jednotlivých binů v rámci buňky rozděleného obrazu. Pro každý pixel v buňce je vypočítána váha, která ovlivňuje histogram orientací hran. Váha je funkce rozsahu gradientu na daném pixelu. Váhy pixelů jsou akumulovány do binů v rámci celého lokálního regionu (buňky). Buňky mohou být obdelníkového nebo oválného tvaru. Biny jsou rovnoměrně rozloženy v rozmezí $0° - 180°$ pro bezznaménkový gradient a v rozmezí $0° - 360°$ pro znaménkový gradient. Váhy pixelů jsou bilineárně interpolovány, aby byl snížen negativní dopad aliasingu. Dalal a Triggs zjistili, že bezznaménkové gradienty v kombinaci s histogramem o devíti kanálech (binech) dosahují nejlepších výsledků v experimentech detekce lidí v obraze. V rámci bloků jsou pak gradienty normalizovány. Výsledný HOG deskriptor je pak vektor skládající se z nor-



Obrázek 3.1: Zjednodušená ilustrace postupu při výpočtu histogramů orientovaných gradientů. Nejprve může být vstupní obraz předzpracován barevnou a gama normalizací. Dále jsou vypočítány gradienty a váhy v jednotlivých buňkách mřížky. V rámci bloků je pak provedena kontrastní normalizace histogramů. Vypočítaná data jsou následně převedena do podoby vhodné pro SVM klasifikátor. Obrázky převzaty z [7].



Obrázek 3.2: Schéma dvou typů kruhových C-HOG bloků. Vlevo je kruhový C-HOG blok s jednotnou kruhovou centrální buňkou a vpravo je ukázka bloku s úhlově rozdělenou centrální buňkou. Obrázky převzaty z [8].

malizovaných histogramů vypočítaných ze všech buněk jednotlivých bloků obrazu. Bloky jsou typicky překrývány přes sebe. Tím je dosaženo různých normalizací v rámci stejných buněk. I přes možnou redundanci dat zajišťuje přesah bloků zlepšení kvality deskriptoru při použití správné normalizace bloků.

Existují dva hlavní typy bloků: obdelníkový R-HOG a kruhový C-HOG. Obdelníkové R-HOG bloky jsou obecně čtvercové mřížky mající několik parametrů: počet buněk v rámci jednoho bloku, rozměry jedné buňky v obrazových bodech a počet kanálů (binů) histogramu vypočítaného v jedné buňce. R-HOG bloky jsou podobné principu SIFT (scale-invariant feature transform) deskriptoru. Ve srovnání se SIFT jsou bloky vypočítány v hustých mřížkách, v určitém měřítku a bez zarovnání orientace, zatímco SIFT deskriptory jsou vypočítávány v řídkých mřížkách s nestejným měřítkem v klíčových bodech obrazu a jsou otáčeny za účelem zarovnání orientace. R-HOG bloky se používají k získání prostorových informací daného objektu, zatímco SIFT deskriptory jsou použity jednotlivě. Kruhové C-HOG bloky existují ve dvou variantách podle jejich geometrie (obrázek 3.2): bloky s jednotnou kruhovou

centrální buňkou a bloky s úhlově rozdělenou centrální buňkou. C-HOG má čtyři hlavní parametry: počet úhlových a radiálních binů, poloměr centrálního binu v obrazových bodech a expanzní faktor pro následné poloměry. Nejméně dva radiální biny (centrální a okolní) a čtyři úhlové biny jsou potřebné k dosažení dobrých výsledků.

HOG deskriptor byl od svého vzniku mnohokrát modifikován a rozšiřován. Například Goto *aj.* [14] vytvořili CS-HOG (color similarity-based HOG) deskriptor založený na segmentaci obrazu podle barevné podobnosti a následném výpočtu HOG příznaků.

3.2 Barevný histogram

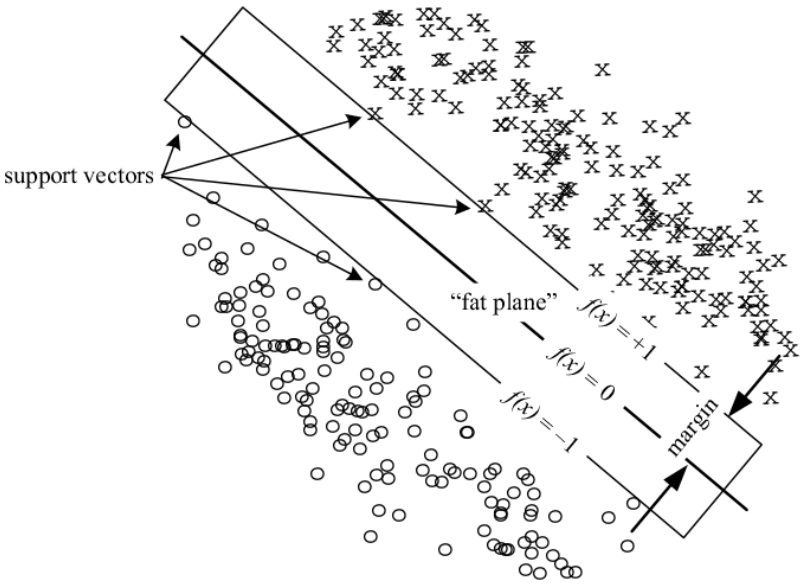
Barevný histogram je jednou z nejzákladnějších metod pro charakterizaci vlastností obrazu. Slouží k reprezentaci rozložení barev v obraze. Do každého barevného rozsahu spadá určitý počet obrazových bodů. Vykreslením těchto rozsahů a počtů vzniká samotný histogram. Barevné histogramy se používají pro různé barevné modely. Nejčastější jsou trojrozměrné modely jako RGB (Red-Green-Blue) a HSV (Hue-Saturation-Lightness) nebo například CMYK (Cyan-Magenta-Yellow-Key). Pro monochromatické obrazy se používá spíše termín histogram intenzit. Pokud je množina možných barev malá, uvádí se jako rozsah barva celá. Potom vzniká histogram, ve kterém každá barva obsahuje určitý počet obrazových bodů. Nejčastější barevné histogramy jsou tvořeny rozsahy v rámci dané barvy.

Pro vytvoření barevného histogramu je nejdříve potřeba rozdělit jednotlivé barvy na počet binů. Každý bin reprezentuje rozsah hodnot dané barvy a počet obrazových bodů, které do tohoto rozsahu spadají. Jedna z výhod barevných histogramů (zejména v oblasti počítačového vidění) je skutečnost, že se podoba histogramu téměř nemění v závislosti na rotaci obrazu. Malé změny je možné zaznamenat při změně úhlu pohledu [25]. Díky tomu je možné srovnávat objekty na základě porovnávání histogramů dvou obrazů bez nutnosti jeho rotace a zarovnání podle hran přes sebe. Barevný histogram tedy nezaznamenává pozici jednotlivých barev, ale jeho hodnoty vychází ze statistik barevného rozložení. Pro účely porovnávání objektů jsou informace pouze v podobě barevných histogramů nedostačující, jelikož mohou nastat situace, kdy naprosto vzhledově rozdílné objekty mají téměř shodné barevné histogramy. Barevný histogram totiž nezaznamenává informace o tvarech nacházející se v obraze. Další nevýhodou je vysoká citlivost na změny osvětlení. Jedny z hlavních výhod barevných histogramů v oblasti počítačového vidění jsou jeho jednoduchý a rychlý výpočet.

Existuje několik metod porovnávání histogramů. Mezi nepoužívanější patří Euklidovská vzdálenost, korelace, chí-kvadrát, průnik a Bhattacharyyova vzdálenost.

3.3 Support vector machines

SVM (support vector machines) je metoda strojového učení. Její původ je v teorii statistického učení, kterou v roce 1998 uvedl Vladimir Vapnik ve svém díle *Statistical learning theory* [28]. Princip metody je hledání roviny v prostoru příznaků, která v optimálním případě rozděluje trénovací data na dva disjunktní poloprostory (obrázek 3.3). Čím více jsou tyto prostory od sebe vzdáleny, tím lépe. Pro učení využívají algoritmy, které analyzují vstupní data a rozpoznávají v nich vzory. SVM se používají pro úkoly klasifikace a regrese. Klasifikace je založena na principu rozhodovaní, do které z již existujících kategorií bude zařazen nový prvek. Při trénování SVM klasifikátoru jsou tedy trénovací data shlukována do kategorií na základě konkrétního měření. Problém regrese spočívá v nalezení vztahu



Obrázek 3.3: SVM v ideálním případě lineárně rozdělitelných dat. Prostor mezi daty definovaný jako $-1 \leq f(x) \leq 1$ je vybrán tak, aby byl co největší. Body nacházející se na hranici tohoto prostoru se nazývají podpůrné vektory (support vectors). Obrázek převzat z [23].

mezi proměnnými. Nejedná se tedy o kategorizovanou odpověď, ale o vyjádření vztahu, ze kterého mohou být předpovězeny případné závěry. V mnohých komplikovaných úlohách nahradily SVM neuronové sítě, které se používaly k řešení zejména klasifikačních problémů. Zároveň je jejich implementace ve srovnání s neuronovým sítěmi podstatně jednodušší [23].

SVM mohou být konceptuálně chápány jako série zobecňování vycházející z ideálního, částečně nereálného počátečního předpokladu. Počáteční předpoklad je speciální případ lineárně rozdělitelných dat. V tomto případě je vycházeno z faktu, že existuje nadrovina v n dimenzích. 1D rovina je definována rovnicí:

$$f(\mathbf{x}) \equiv \mathbf{w} \cdot \mathbf{x} + b = 0$$

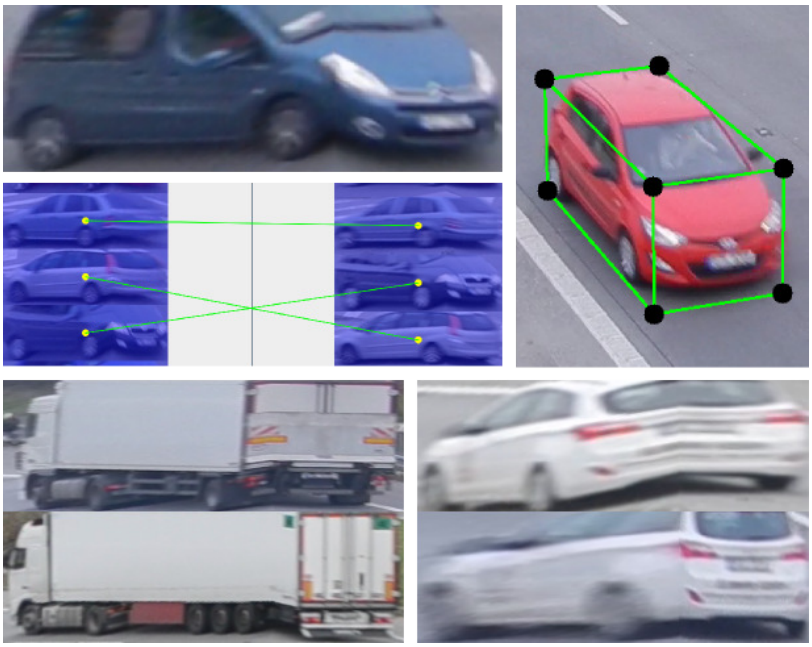
Tato rovina kompletně odděluje trénovací data. Respektive všechny trénovací instance označeny jako pozitivní ($y_i = 1$) leží na jedné straně nadroviny ($f(\mathbf{x}_i) > 0$), zatímco všechny instance označeny jako negativní ($y_i = -1$) leží na druhé straně nadroviny ($f(\mathbf{x}_i) < 0$). Proto stačí nalézt normálový vektor k nadrovině $\mathbf{w}$ a ofset b. Potom se v rovnici stane $f(\mathbf{x})$ rozhodovacím pravidlem. Obecně může existovat více jak jedna nadrovina, která lineárně rozděluje data. Data nacházející se v těsné blízkosti hranice nadroviny se nazývají podpůrné vektory. V případě problému klasifikace nesou tyto vektory všechny relevantní informace [17]. Spolu s hledáním uvedeného rozhodovacího pravidla je řešen i problém kvadratického programování. Dále jsou u SVM často využívány tzv. jádra (kernel), které zobecňují lineární algoritmus na nelineární. Použití nelineárních SVM umožňuje řešení komplexnějších a složitějších problémů než v případě lineárních. Výpočetní náročnost ovšem s jejich použitím rapidně stoupá. Proto je vhodné, pokud to je možné, řešený problém přeformulovat tak, aby byl lineárně řešitelný. Regrese u SVM vychází ze ztrátových funkcí. Ztrátová funkce ignoruje chybovost, která je menší než stanovený práh ϵ. Přesnost regrese může být odhadnuta pomocí Laplaceovy distribuce.

Nejen v počítačovém vidění hrají SVM velice významnou roli. Pomocí strojového učení je možné řešit různé netriviální problémy, které vyžadují komplexnější znalosti. Jedním z typických problémů, na kterém se demonstrují principy a funkčnost SVM je nalezení nevyžádané pošty (spam) v příchozích e-mail zprávách. SVM klasifikátoru jsou předloženy vhodně reprezentované příklady nevyžádané pošty a obyčejných zpráv. Podle výskytů jednotlivých druhů slov a jazykových formulací dokáže SVM klasifikátor rozhodnout, zda-li se jedná o nevyžádanou poštu či nikoliv.

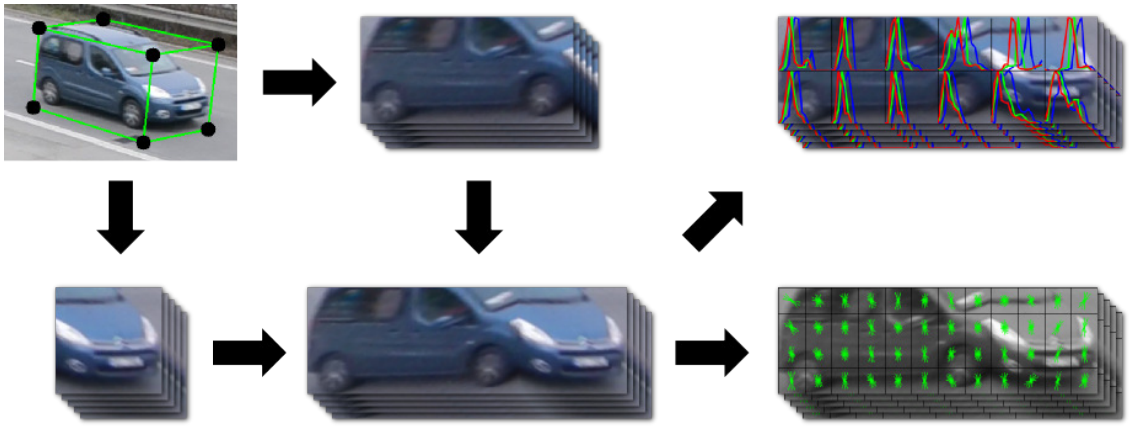
Kapitola 4

Návrh a implementace systému pro reidentifikaci automobilů ve videu

Návrh a řešení problému reidentifikace automobilů ve videu (obrázek 4.1) popsané v této kapitole je založeno na předpokladu, že detekce vozidla v obraze je již vyřešena. Detekční systém, který jsem pro reidentifikaci použil, se automaticky zkalibruje na danou scénu



Obrázek 4.1: Tato práce řeší problém reidentifikace automobilů zaznamenaných různými kamerami. Vstup do systému jsou detekované automobily v obraze. Následně jsou extrahována data z 3D ohraničujících kvádrů vytvořených pomocí plně automatické kalibrace vycházející z dané scény [9, 10] *(vpravo nahoře)*. Obrazová data jsou z 3D ohraničujícího kvádru promítnuta do roviny za účelem koncentrace pouze na vozidlo samotné s minimálním množstvím nechtěných informací *(vlevo nahoře)*. Na základě manuálních anotací *(vlevo uprostřed)* je vytvořena trénovací datová sada. **dole:** Ukázka typického páru vyhodnoceného systémem jako pozitivní a páru velice podobných vozidel správně vyhodnocených jako rozdílné.



Obrázek 4.2: Zjednodušená ilustrace extrakce dat a jejich reprezentace pomocí barevného histogramu a histogramu orientovaných gradientů. Nejdříve je automobil detekován ve videu. Výsledek detekce je reprezentován jako ohraničující kvádr (bounding box) kolem automobilu. Následně jsou vyextrahovány a promítnuty do roviny (rektifikovány) boční a přední stěny všech ohraničujících kvádrů náležící k danému automobilu. Dále jsou k sobě náležící páry rektifikovaných bočních a předních stěn spojeny k sobě. Takto vznikne množina několika sloučených snímků zachycující přední a boční stranu automobilu. Deformace spojeného snímku je způsobena perspektivou zachycené scény. Velikost množiny záleží na vlastnostech videa (druh scény, úhel záběru, počet snímků za sekundu,...) a přesnosti detekce automobilu ve videu. Nakonec jsou z daných spojených snímků vypočítány informace o barvě v podobě barevných histogramů a informace o tvaru v podobě histogramů orientovaných gradientů. Vykreslení histogramů na obrázku je pouze ilustrativní.

a následně vytváří 3D ohraničující kvádry kolem detekovaných vozidel, se kterými se dále pracuje [9, 10].

Samotná reidentifikace je založena na regresi, která je vyhodnocena lineárním klasifikátorem a která vychází z dvou natrénovaných modelů: model používající barevné histogramy [26] a model založený na histogramu orientovaných gradientů (HOG) [8]. Před natrénováním lineárního klasifikátoru je nejdřív potřeba extrahovat potřebné informace o vozidle z ohraničujících kvádrů. Proto jsou nejdříve promítnuty boční a přední stěny ohraničujících kvádrů do roviny a následně jsou spojeny (kapitola 4.2). Vzniká tak tzv. kombinovaný obraz, který je pro extrakci dostatečného množství informací rozdělen do mřížky o příslušné velikosti. Pro každé okénko mřížky jsou pak vypočítány hodnoty barevného histogramu a hodnoty histogramu orientovaných gradientů (kapitola 4.3). Dosavadně zmíněný postup je ilustrován obrázkem 4.2.

Na základě vypočítaných hodnot jsou pak vytvořeny oba modely trénováním lineárního klasifikátoru na dostatečném množství pozitivních a negativních vzorků (kapitola 4.4). Konečná reidentifikace vychází z výsledků regresí vypočítaných klasifikátorem a z nalezení nejvhodnějšího automobilu v databázi detekovaných automobilů z jiných kamer na základě předešle získaných dat (kapitola 4.5).

Přesnost systému byla vyhodnocena různými způsoby. Mimo trénovací sadu pro lineární klasifikátor byly ve stejném čase pořízeny záznamy různými kamerami, natočeny z různých úhlů a různě přiblíženy na stejnou dálnici. Data byla manuálně oanotována a systém vyhodnocen na nich. Jelikož mohou nastávat i případy, kdy se automobil vzhledově shoduje

s jiným automobilem v databázi, ale nejedná se o týž automobil[1], byla vybrána podmnožina podobných automobilů a vytvořeno webové rozhraní, pomocí kterého mohli lidé rozhodovat, zda-li „by se mohlo jednat o stejný automobil". Díky způsobu vyhodnocení zmíněném v kapitole 5 bylo možné rozdělit datovou sadu na množinu případů jednoznačných shod (automobily jsou vzhledově stejné), neshod (automobily jsou zcela rozdílné) a množinu případů, kdy se lidé nebyli schopni úplně shodnout (u automobilů bylo z různých důvodů velice těžké rozlišit jejich totožnost).

Uvedené řešení jsem implementoval v programovacím jazyce C++ s využitím knihovny OpenCV (opensource computer vision) [5], která nabízí mnoho knihovních funkcí pro zpracování obrazu v reálném čase. Grafické uživatelské rozhraní pro vytváření párů stejných automobilů uvedené v kapitole 4.4 jsem naprogramoval v jazyce Java. Pro trénování lineárního klasifikátoru a výpočet regrese byla použita knihovna LIBLINEAR [11].

4.1 Základní obrysy návrhu řešeného systému

Na základě poznatků získaných v kapitolách 2 a 3 mohu nyní vytvořit návrh systému pro reidentifikaci automobilů ve videu.

Detekce automobilů ve videu je již vyřešena, tudíž bude systém orientován jen na reidentifikaci samotnou. Je ovšem nutné informaci o detekci automobilu vhodně zpracovat. Detekční systém generuje informace o automobilu v následujícím tvaru:

```
1,1816,648,240,521,243,405,183,500,181,652,415,528,423,412,316,506,312
```

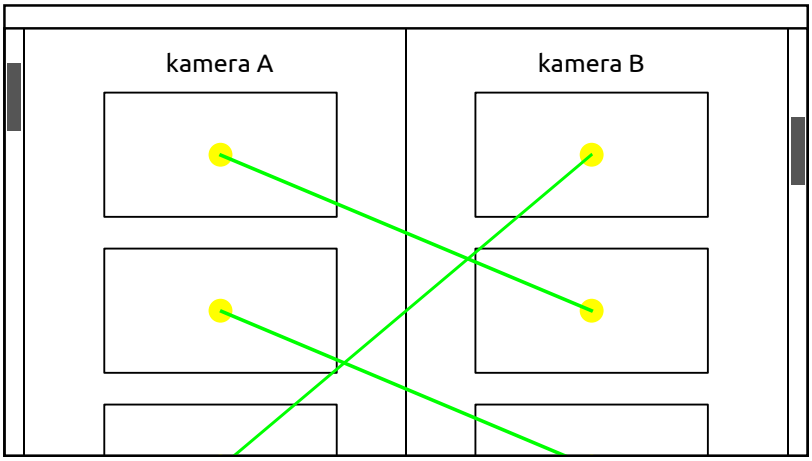
První číslo je identifikátor automobilu. Druhé číslo je snímek ve video záznamu, na kterém se daný automobil nachází. Za ním následují souřadnice jednotlivých bodů 3D ohraničujícího kvádru, který slouží jako prostorová reprezentace detekce automobilu v obraze. Tyto informace budou sloužit jako vstup do systému, proto je bude potřeba rychle a správně zpracovávat.

Po zpracování vstupních informací o detekovaném vozidle bude potřeba automobily vyextrahovat z obrazu. Díky použitému 3D ohraničujícímu kvádru v obraze je možné jednoduše pracovat s jeho jednotlivými stěnami. Extrahované automobily bude potřeba uložit, jelikož budou sloužit pro trénování lineárního klasifikátoru a také pro experimentování s různými metodami pro porovnávání a extrahování příznaků.

Pro jednoduché a rychlé vytvoření vazeb mezi automobily za účelem trénovaní jsem navrhl jednoduchý vzhled aplikace (obrázek 4.3), která umožní efektivně vytvořit páry stejných automobilů ve velké datové sadě. Informace vytvořené touto aplikací bude systém používat pro načítání vozidel do paměti a také pro trénování lineárního klasifikátoru.

Pro dosažení maximální přesnosti a úspěšnosti reidentifikace automobilů je nutné v systému vhodně reprezentovat klíčové vizuální prvky daného vozidla. Zároveň je žádoucí, aby data reprezentující vzhled automobilu obsahovala co nejmenší množství informací, které se v obraze vozidla netýkají. Jedná se zejména o maximální redukci obrazových pixelů náležící prostoru za vozidlem. Pozadí automobilu může totiž negativně ovlivňovat výslednou reidentifikaci. Mnohem obtížnější úkol je redukce obrazových informací objektů, se kterými se v záběru vozidlo dostává do kolize (keř, pouliční lampa, jiný automobil, ...). Tato data budou sloužit spolu s vazbami vytvořenými zmíněnou aplikací jako vstupní informace pro

[1] Z hlediska reidentifikace samotné se jedná o falešně pozitivní vyhodnocení, které ovšem nevypovídá o přesnosti systému. Pro věrohodné vyhodnocení systému je proto potřeba zahrnout i toto malé procento případů



Obrázek 4.3: Návrh vzhledu aplikace pro anotování stejných párů automobilů, které budou sloužit jako výchozí data pro trénování lineárního klasifikátoru. Čáry na obrázku znázorňují vazby mezi obrázky jednotlivých vozidel.

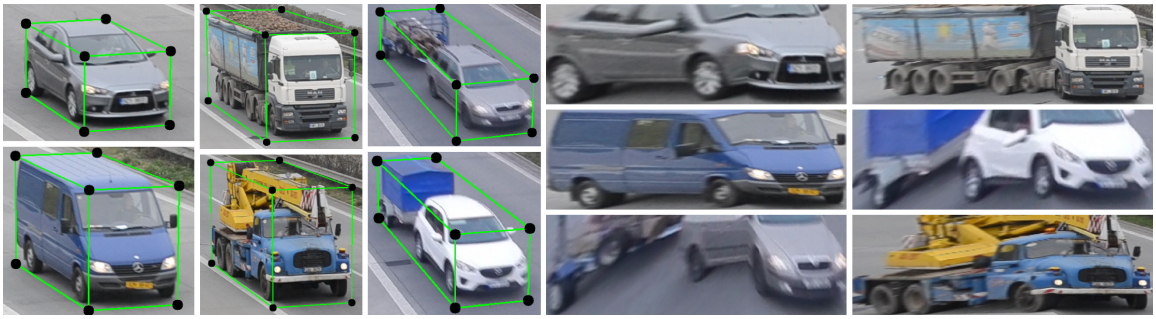
natrénování lineárního klasifikátoru, který bude vyhodnocovat podobnost jednotlivých párů automobilů. V neposlední řadě je důležité data o vozidle extrahovaná z obrazu uskladnit a zajistit rychlý přístup k nim. Urychlí se tak vyhledávání v databázi automobilů a z toho vyplývající počet možných reidentifikací za sekundu.

Po natrénování klasifikátoru je důležité nalézt co nejrychlejší způsob vyhledávání hledaného automobilu v databázi, aby bylo možné reidentifikaci provádět v reálném čase.

4.2 Detekce a extrakce automobilu z obrazu

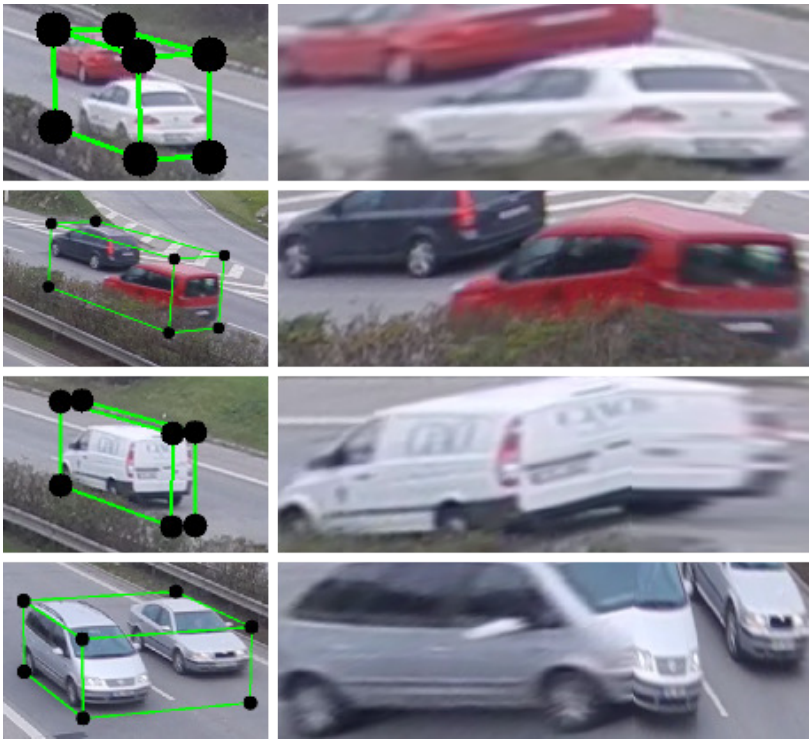
Aby došlo k redukci nežádoucích dat extrahovaných z obrazu a zároveň aby byly zachovány informace o proporcích automobilu (přední/zadní maska, boční strana, ...), je v detekčním systému automobilů použit 3D ohraničující kvádr jako reprezentace detekce (obrázek 4.4). Díky 3D ohraničujícímu kvádru je možno pracovat s jednotlivými stěnami kvádru zvlášť. Výhodou tohoto přístupu je jednoduchá práce s daty v obraze a již zmíněná značná redukce „šumu" v pozadí automobilu bez nutnosti dalšího zpracování obrazu a z toho vyplývajícího snížení výpočetního výkonu, které by bylo nutné pro odstranění pozadí v případě použití jednoduchého 2D hraničního obdelníku.

Použitý detekční systém vytváří vstup pro reidentifikaci v podobě vektoru o sedmi bodech (3D ohraničující kvádr). Spolu s nimi je obdrženo jedinečné identifikační číslo automobilu v rámci daného záběru a číslo snímku, na kterém se automobil vyskytuje. Po zpracování a uložení těchto informací následuje vyjmutí automobilu z obrazu na daném snímku. Za účelem optimalizace a redukce datové náročnosti algoritmu jsem se rozhodl použít pouze boční a přední stěnu ohraničujícího kvádru, jelikož jsem studiem datové sady došel k závěru, že střecha automobilu v převážné většině případů nenese žádnou užitečnou informaci. Následně jsou dané strany ohraničujícího kvádru promítnuty do roviny a spojeny v jeden obraz (obrázek 4.4, tři řádky vpravo). V průměru obsahuje tímto způsobem vyextrahovaný obraz přibližně 5 % plochy nežádoucích informací (pozadí), což je v porovnání s 25 % při použití jednoduchého hraničního obdelníku značný rozdíl. Obrázek 4.5 ukazuje rozdíl mezi zmíněnými případy. Výsledný rektifikovaný obraz je viditelně zdeformován z důvodu perspektivy dané scény, což pro účely reidentifikace nevadí.



Obrázek 4.4: Příklady správné detekce vozidla v obrazu a vykreslení ohraničujícího kvádru. Vpravo jsou zobrazeny odpovídající extrahovaná vozidla z ohraničujícího kvádru včetně rektifikace a spojení jeho boční a přední stěny. Vstupní data získaná z detekčního systému jsou reprezentována jako vektor 7 bodů popisující vrcholy ohraničujícího kvádru.

Obrázek 4.5: Srovnání extrakce automobilu z obrazu za použití jednoduchého 2D hraničního obdelníku (nahoře) a 3D ohraničující kvádru (dole) pro detekci automobilu. Rektifikace a spojení boční a přední strany ohraničujícího kvádru způsobí deformaci výsledného snímku. Tímto způsobem je dosaženo značné redukce nechtěných dat nacházejících se za automobilem. Oproti původním 25 % tvoří pozadí snímku pouhých 5 % plochy. Proto je pro zpřesnění výsledků reidentifikace výhodnější použít ohraničujícího kvádr jakožto reprezentace detekce automobilu ve videu.

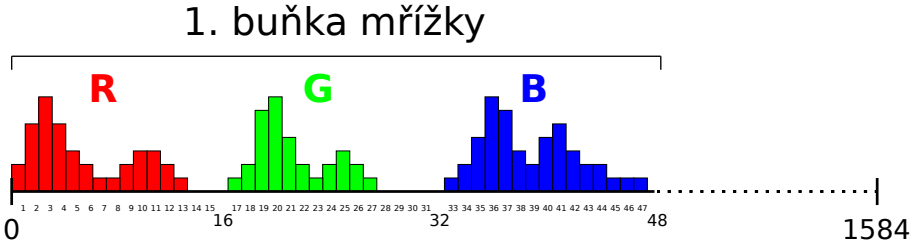


Obrázek 4.6: Příklady nesprávné detekce vozidla v obrazu včetně extrakce a rektifikace z ohraničujícího kvádru. Nepřesnost detekce vzniká zejména v případech, kdy jeden automobil předjíždí druhý a jsou pak detekovány jako jeden automobil nebo se v obraze automobil dostává do kolize s jiným objektem (například keř). V některých případech, například u automobilů s přívěsem, trvá přibližně 3−4 snímky videa, než se detekční systém zkalibruje a ohraničující kvádr umístí po celém rozměru soupravy.

Přesnost detekčního systému není naprosto dokonalá a přibližně v 5 % detekcí selhává. Míra chybovosti záleží na více faktorech: hustota provozu, pohled kamery, viditelnost, atd. Obrázek 4.6 ukazuje několik druhů chyb detekčního systému. Po promítnutí chybných detekcí do roviny pak vzniká 2D reprezentace obrazu, která neodpovídá vzhledu daného automobilu. Avšak kromě poměrně malé extra zátěže procesoru způsobené zpracováním chybné detekce nepředstavuje tato skutečnost značnou újmu na přesnosti výsledné reidentifikace, jelikož jsou případy chybné detekce v celé sadě snímků, na kterých se konkrétní automobil vyskytuje, ojedinělé.

4.3 Příznaky automobilu

V této kapitole jsou popsány principy extrakce informací o vozidle po jeho úspěšné detekci, vyjmutí z obrazu a rektifikaci. Příznaky automobilů jsem se rozhodl reprezentovat v podobě sady barevných histogramů a histogramu orientovaných gradientů. Pro urychlení reidentifikace je pak z každé ze sad vyextrahovaných příznaků (cca 25 na jedno vozidlo) vypočítána i průměrná informace vycházející z uvedených metod extrakce příznaků. Tato informace je použita pro předvýběr automobilů, mezi kterými následně probíhá výpočetně náročnější hluboká regrese. V následujících kapitolách popíši podrobnosti a parametry zmíněných metod, které v průběžném experimentování dosahovaly nejlepšího poměru rychlosti a přesnosti reidentifikace.

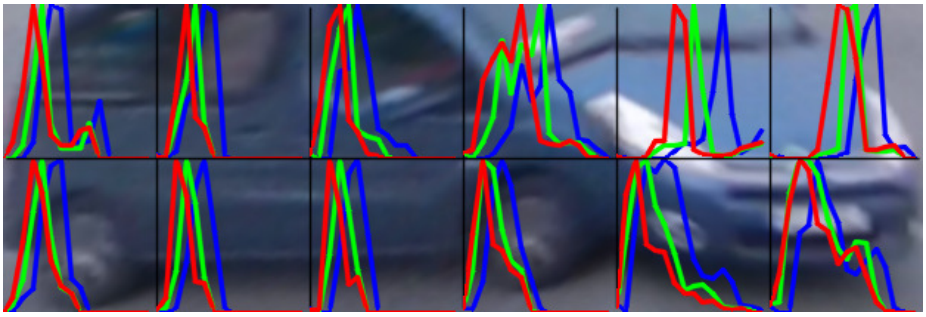


Obrázek 4.7: Výsledek reprezentace dat získaných z jednoho snímku tvořeného spojenou rektifikovanou boční a přední stěnou 3D ohraničujícího kvádru detekovaného automobilu. Reprezentace je v podobě jednoho vektoru skládajícího se z jednotlivých hodnot všech tří složek barevných histogramů vypočítaných v rámci všech buněk mřížky. V jedné buňce je celkově vypočítáno 48 hodnot (16-R, 16-G, 16-B). Celkově vektor obsahuje 1548 hodnot. Počet hodnot ve výsledném vektoru lze vypočítat podle vzorce $\left(\frac{x_n-1}{x_s}+1\right)\cdot\left(\frac{y_n-1}{y_s}+1\right)\cdot b\cdot c$. Hodnota x_n reprezentuje počet buněk mřížky v ose x. Hodnota x_s vyjadřuje přesah buněk ve směru osy x. Přesah buněk musí být hodnota v intervalu $(0, 1\rangle$. Analogicky pro hodnoty y. Hodnota b vyjadřuje počet sloupců jedné barevné složky barevného histogramu a hodnota c pak počet barevných složek histogramu. Obdobně je prováděna reprezentace dat vycházející z histogramu orientovaných gradientů.

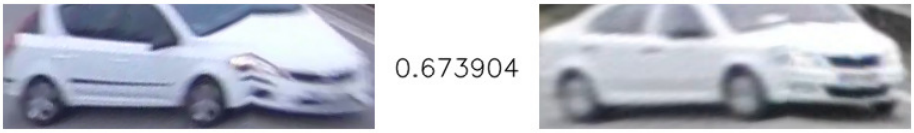
4.3.1 Barevný histogram

Některé z uvedených prací v kapitole 2 nebraly v potaz informaci o barevných vlastnostech vozidla. V některých případech to může být rozumné, zejména za účelem redukce extrahovaných dat o vozidle, které je nutné pro reidentifikaci uschovávat. Pro dosažení maximální úspěšnosti anonymní reidentifikace, jejíž hlavním zdrojem informací je právě vzhled daného objektu, jsem se rozhodl brát barvu vozidla v potaz. Nejen že informace o barevných vlastnostech automobilu výrazně ovlivňuje výpočet hodnoty regrese a snižuje množství falešně pozitivních reidentifikací, ale zároveň výrazně zrychluje reidentifikaci samotnou, jelikož je brána jako první rozhodující kritérium pro podobnost automobilů. Výpočet barevných hodnot obrazu také patří mezi jednu z nejrychlejších metod výpočtu příznaků [26]. Pokud jsou tedy porovnávány evidentně barevně rozdílné automobily, zamezí se tak zbytečným dalším výpočtům.

Pro reprezentaci barevné informace o vozidle byl použit barevný histogram pro každou z RGB složek obrazu. Přesněji se jedná o barevný histogram o 16-ti sloupcích. Histogram o více sloupcích téměř vůbec nezlepšoval úspěšnost reidentifikace a razantně zvyšoval datovou náročnost. Naopak histogram o méně sloupcích sice zrychloval jeho samotný výpočet a nebyl tak datově náročný, avšak vypočítaná data nedostačovala pro jednoznačné rozhodnutí o barevné podobnosti porovnávaných automobilů. Spojený obraz přední a boční rektifikované stěny ohraničujícího kvádru je rozdělen do mřížky o velikost 6×2. Pro každou buňku mřížky je pak vypočítán barevný RGB histogram. Aby došlo k maximálnímu využití dostupných barevných informací v obraze, je vypočítán následující barevný histogram s 50 % přesahem velikosti jedné buňky mřížky v obou směrech osy x a y. Z toho vyplývá, že výsledný rozměr mřížky lze teoreticky považovat za 11×3. Rozměr 6×2 tedy definuje velikost posunujícího se okénka. Hodnoty histogramu jsou průběžně ukládány do jednoho vektoru (obrázek 4.7). Celkově je v rámci jednoho spojeného obrazu vozidla vypočítáno 1 584 hodnot. Počet spojených obrazů náležících jednomu vozidlu záleží na přesnosti detekce vozidla, vlastnostech scény, počtu snímků za sekundu daného záznamu, atd. V průměru se



Obrázek 4.8: Vykreslení barevných histogramů spojeného obrazu automobilu. Obraz je rozdělen do mřížky 6×2. Vykreslení je pro přehlednost zobrazeno bez přesahů buněk mřížky (12 histogramů oproti původním 33).

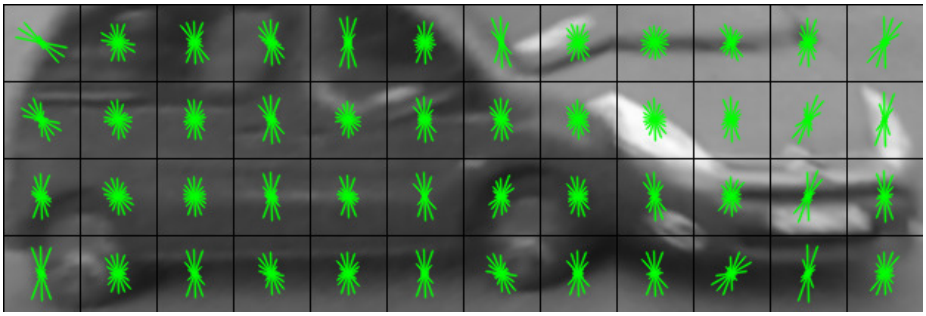


Obrázek 4.9: Výsledek reidentifikace založené pouze na barevných histogramech. Mezi obrázky automobilů je výsledná hodnota regrese (viz kapitola 4.5). Algoritmem byla získána poměrně vysoká hodnota (možno srovnat s obrázkem 5.6). I přes podobné vizuální charakteristiky vozidel se nejedná o vozidla totožná. Z toho vyplývá, že informace pouze o barvě vozidla nejsou dostačující k dosažení úspěšné reidentifikace.

jedná přibližně o 25 spojených snímků náležící jednomu vozidlu. Na obrázku 4.8 je ukázka vykreslených barevných RGB histogramů bez přesahu buněk mřížky na jednom spojeném obrazu automobilu.

4.3.2 Histogram orientovaných gradientů

Pro dosažení úspěšné reidentifikace automobilu je nutné získat více informací z obrazu, jelikož pouze informace reprezentované barevnými histogramy nejsou dostačující (i přes jejich rychlý výpočet a eliminaci mnoha kandidátů pouze na základě barevných informací). Výsledek reidentifikace založený pouze na barevných RGB histogramech je vidět na obrázku 4.9. Proto jsou spolu s barevnými histogramy vypočítány i histogramy orientovaných gradientů (HOG) [26], které vypovídají o tvaru automobilu. HOG informace je uložena odděleně od barevných histogramů. Histogram orientovaných gradientů je vypočítán pro všechny spojené obrazy náležící jednomu automobilu. Obraz je nejdříve převeden do odstínů šedi a následně rozdělen do mřížky o rozměru 12×6. Při výpočtu histogramů orientovaných gradientů dochází stejně jako u RGB histogramů k 50 % přesahu buněk. Z toho vyplývá, že je celkově vypočítáno 161 histogramů orientovaných gradientů z jednoho spojeného obrazu automobilu. Jeden histogram má 9 sloupců. Výsledný vektor, do kterého jsou hodnoty průběžně ukládány (obdobně jako na obrázku 4.7), má dohromady 1 449 položek. Na obrázku 4.10 je ukázka vykreslených histogramů orientovaných gradientů bez přesahu buněk uvedené mřížky.



Obrázek 4.10: Vykreslení histogramů orientovaných gradientů spojeného obrazu automobilu. Obraz je rozdělen do mřížky 12×4. Vykreslení je pro přehlednost zobrazeno bez přesahů buněk mřížky (48 histogramů oproti původním 161). Pro výpočet histogramu orientovaných gradientů je nutno převést vstupní obraz do odstínů šedi.

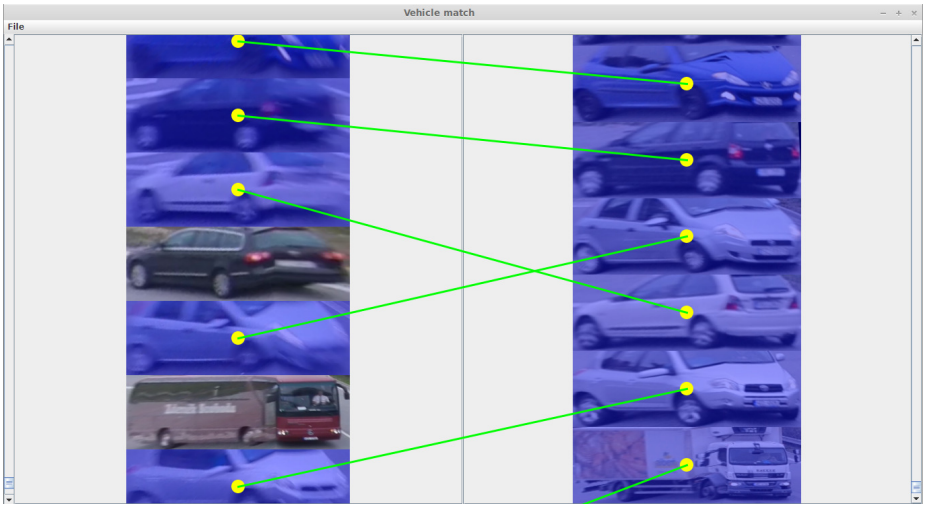
4.4 Trénování lineárního klasifikátoru

Rozhodování o podobnosti/shodě dvou automobilů vychází z výsledků získaných z lineárního SVM klasifikátoru [11]. Jedna z hlavních výhod tohoto typu klasifikátoru je, že může být poměrně rychle natrénován na velkém množství trénovacích vzorků. Stejně tak může jeden vzorek obsahovat mnoho hodnot. Regrese je pomocí lineárního klasifikátoru ve srovnání s nelineárním SVM klasifikátorem také mnohem rychlejší.

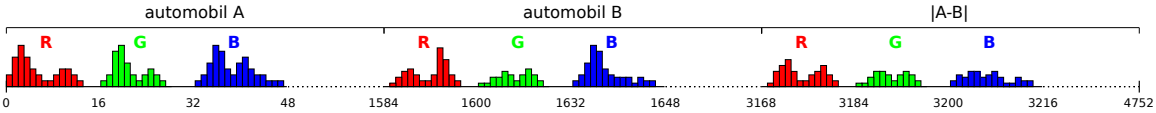
Pro přesné výsledky regrese je potřeba klasifikátor natrénovat na dostatečném množství pozitivních a negativních vzorků. Proto bylo pořízeno celkově pět záběrů snímající dálnici. Záběry byly natočeny ze stejného místa avšak z jiného úhlu, různými typy kamer a různě transfokovány. Tím bylo dosaženo simulace různých prostředí, ve kterých se mohou kamery nacházet, a zároveň zjednodušeno následné anotování vozidel díky časové synchronizaci záběrů. Celkově bylo natočeno přibližně 800 automobilů.

Za účelem usnadnění anotací dat pro trénování lineárního klasifikátoru bylo vytvořeno jednoduché grafické uživatelské rozhraní (obrázek 4.11). Pomocí vytvořeného GUI lze rychle spárovat větší množství automobilů vyextrahovaných z natočených záběrů. Data získaná z GUI aplikace pak slouží jako výchozí informace pro trénování.

Vzhled a tvar vozidel je reprezentován pomocí vektorů barevných histogramů a histogramů orientovaných gradientů vypočítaných z množiny vyextrahovaných obrazů náležící danému automobilu a pomocí z nich vypočítaných průměrných vektorů, jak již bylo zmíněno v kapitole 4.3. Trénovací data vychází z těchto vektorů. Respektive jsou spojeny (konkatenovány) dohromady včetně jednoho rozdílového vektoru (obrázek 4.12). Spojené vektory jsou vytvořeny ze stejných a různých dvojic automobilů. Konkatenované vektory stejných automobilů jsou vytvořeny z náhodných kombinací příznakových vektorů vypočítaných z množin obrazů stejných automobilů (příklad na obrázku 4.13). Při trénování lineárního klasifikátoru bylo použito přibližně $12\,000-16\,000$ pozitivních vzorků. Konkatenované vektory různých automobilů jsou vytvořeny z dvojic náhodně vybraných automobilů natočených kamerou A a B. Pravděpodobnost vytvoření negativního trénovacího vzorku z dvojice stejných vozidel je při velikosti vytvořené datové sady velice malá. Pro zpřesnění regrese bylo vygenerováno a použito pro trénování dvojnásobné množství negativních vzorků oproti vzorkům pozitivním. Výsledkem procesu trénování je model pro příznakové vektory vycházející z histogramů orientovaných gradientů a model pro příznakové vektory vycházející z barevných histogramů.



Obrázek 4.11: Grafické uživatelské rozhraní pro anotování trénovacích dat, pomocí kterého je možné vytvářet dvojice totožných automobilů získaných z různých kamer. Modře vybarvené automobily mají svého dvojníka z jiného zdroje a jsou propojeny zelenou čarou. Nevybarvené automobily svého dvojníka nemají (byl vyřazen z důvodu chybné detekce nebo nebyl detekován vůbec) a zůstávají tudíž neoznačeny. Data získaná z aplikace slouží jako základní zdroj informací pro lineární klasifikátor.



Obrázek 4.12: Ukázka vektoru určeného pro následnou regresi nebo pro natrénování lineárního klasifikátoru. V případě trénování je takový vektor označen hodnotou 1.0 (automobily A a B jsou shodné) nebo hodnotou -1.0 (automobily A a B jsou rozdílné). Při regresi v procesu reidentifikace je pak po vyhodnocení takového vektoru obdržena hodnota blížící se jedné ze zmíněných trénovacích hodnot (tedy svědčí o barevné podobnosti daného páru automobilů). Vektor je tvořen z RGB hodnot barevných histogramů automobilů A a B. Dále je na konec vektoru přidán jejich absolutní rozdíl. Celkově výsledný vektor obsahuje 4 752 hodnot.

4.5 Regrese a nalezení vhodného kandidáta pro reidentifikaci

Samotná reidentifikace se skládá z několika kroků. Nejdříve jsou použity příznakové vektory průměrných hodnot vytvořené ze všech příznakových vektorů skládajících se z hodnot histogramů orientovaných gradientů a hodnot barevných RGB histogramů. Podle principu uvedeném na obrázku 4.12 jsou vytvářeny párové vektory průměrných příznakových vektorů RGB histogramů z množiny automobilů, kde by se mělo reidentifikované vozidlo nacházet, s automobilem, které má být reidentifikováno. Vytvořené vektory jsou vyhodnocovány klasifikátorem jako regrese v podobě desetinného čísla. Pokud je výsledek regrese nad nastaveným prahem, prošel pár prvním kolem regrese a je testován dále. Hodnota prahu je odvozena od požadavků na systém a vypovídá o hodnotě dělícího kritéria (viz kapitola 5). Páry automobilů, které byly označeny za pozitivní na základě průměrných příznakových vektorů barevných RGB histogramů jsou stejným způsobem vyhodnocovány v druhém kole,

Obrázek 4.13: Ukázka náhodně vybraných dvojic stejného vozidla natočeného z různých kamer. Data získaná z takových dvojic slouží jako vstup pro trénovaní lineárního klasifikátoru na stejné automobily. **vlevo**: Kamera A, **vpravo**: Kamera B.

kde se ovšem vychází z průměrných příznakových vektorů histogramů orientovaných gradientů. Automobily vyhodnocené v obou kolech jako pozitivní s hledaným automobilem jsou pak přidány do množiny automobilů, které jsou si s hledaným automobilem vizuálně velice podobné. Například při reidentifikaci automobilu v sadě 300 vozidel obsahuje tato množina přibližně 5−10 automobilů. Velikost množiny závisí na různorodosti vizuálních charakteristik automobilů v dané sadě. Uvedeným způsobem bylo dosaženo efektivní a výpočetně nenáročné (maximálně 2 porovnání na jeden pár) filtrace množiny automobilů na podmnožinu vizuálně podobných vozidel s hledaným automobilem.

Nalezení reidentifikovaného vozidla v uvedené podmnožině je založeno na hluboké regresi s hledaným automobilem. Hluboká regrese spočívá v několika regresích vyhodnocených z náhodně vybraných příznakových vektorů HOG a barevných RGB histogramů náležící porovnávaným automobilům (nejedná se již o průměrný příznakový vektor). Hluboká regrese se může například skládat z dílčích regresí vypočítaných z párů na obrázku 4.13. Složitost hluboké regrese (počet dílčích regresí) se podobně jako práh regrese odvíjí od požadavků na systém (rychlost, úspěšnost, ...). Výsledky dílčích regresí jsou průběžně ukládány a následně je z hodnot vypočítán medián. Automobil, který v porovnávání s hledaným automobilem dosáhl nejvyšší hodnoty hluboké regrese nad stanoveným prahem, je považován za úspěšně reidentifikovaný.

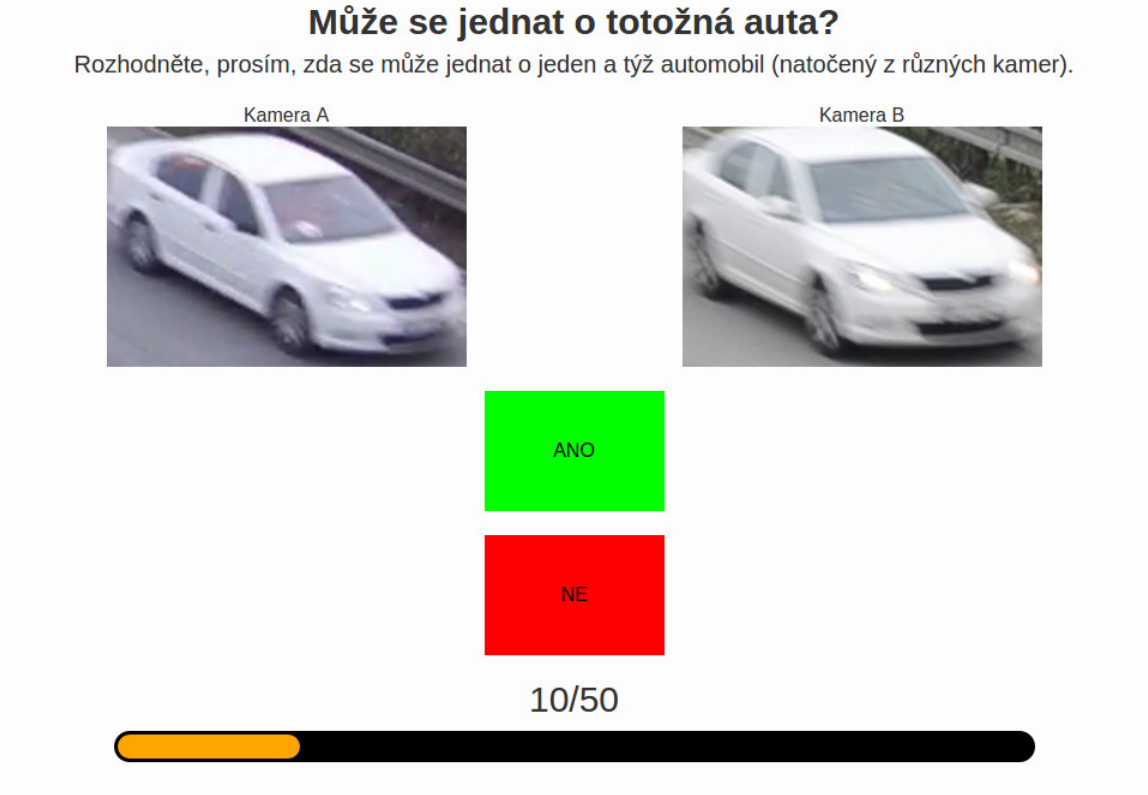
Kapitola 5

Experimenty a dosažené výsledky

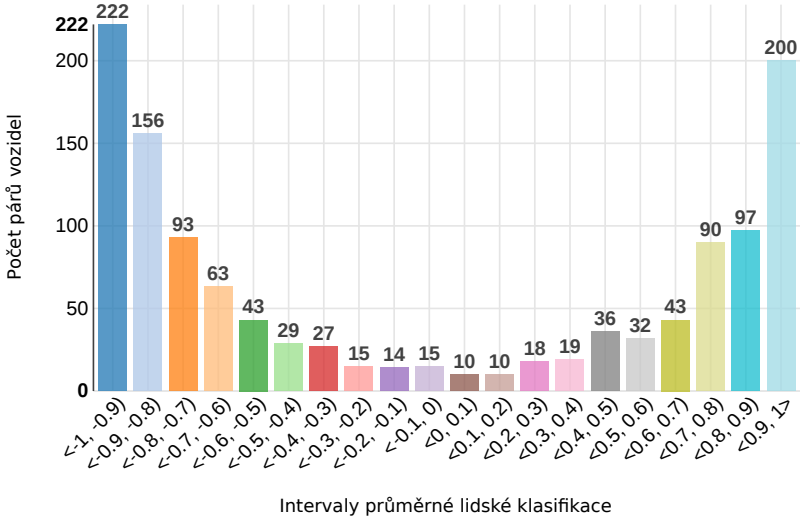
V této kapitole jsou zmíněny způsoby, jakým byl implementovaný algoritmus vyhodnocen. Pro dosažení požadovaných výsledků bylo nutné experimentovat s různými hodnotami parametrů reidentifikačního algoritmu. Dále jsou zde rozebírány případné možné budoucí změny a vylepšení, které by mohly na základě získaných poznatků vést k přesnějším výsledkům algoritmu.

Pro získání představy o vypočítaných hodnotách podobnosti porovnávaných párů automobilů jsem vygeneroval ukázkovou sadu zobrazující výsledky hlubokých regresí z kartézského součinu množiny automobilů A (828 vozidel) a B (720 vozidel), tj. celkem 645 840 kombinací. Část získaných výsledků jsou na obrázcích 5.5 a 5.6 (pro velikost jsou obrázky uvedeny na konci kapitoly). Z obrázků je patrný význam hodnot regresí vypočítaných lineárním klasifikátorem. Zobrazená škála hodnot podobnosti jednotlivých párů byla vypočítána nastavením stejné váhy pro výsledky dílčí regrese vycházející jak z hodnot barevných histogramů tak z hodnot histogramů orientovaných gradientů. Výsledek dílčí regrese získaných z obou modelů byl zprůměrován a zapamatován. Pro jeden pár automobilů bylo získáno 21 zprůměrovaných hodnot dílčích regresí. Z těchto hodnot byl následně získán medián, který je výslednou hodnotou hluboké regrese na zmíněných obrázcích.

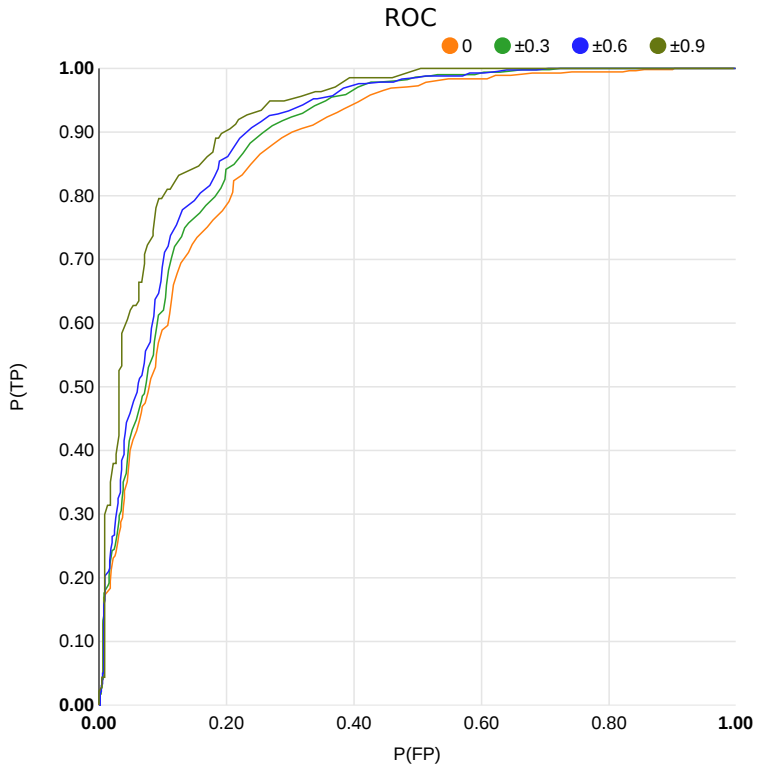
Za účelem kvantitativního vyhodnocení byla ze stejného množství párů algoritmem vybrána množina 1 232 párů automobilů, které si jsou jak barevně tak tvarově podobné a mohou být stejné (práh regrese nebyl striktně nastaven pouze na shodné případy). Pro ohodnocení vybraných párů bylo vytvořeno webové rozhraní (obrázek 5.1), pomocí kterého mohli lidé rozhodovat, zda-li „se může jednat o totožná auta" natočená z různých kamer. Tímto způsobem bylo obdrženo přibližně 24 000 rozhodnutí (klasifikací ohodnocených jako 1 pro pár stejných automobilů a -1 pro pár rozdílných automobilů) od více než 500 lidí, což činí necelých 20 anotací na jeden pár automobilů. Průměrné výsledky anotací jsou znázorněny histogramem na obrázku 5.2. Z grafu lze vyčíst, že na přibližně jedné třetině párů automobilů se lidé jednoznačně shodli (průměrná hodnota z anotací náležící danému páru je rovna nebo se velmi blíží hodnotám 1 a -1) a značná část párů byla ohodnocena jako většinové stejná nebo rozdílná. Avšak nezanedbatelná část párů byla ohodnocena poměrně nejednoznačně (průměrná klasifikace se blíží hodnotě 0). Takové páry se většinou lišily v barevném tónování oken nebo z daného úhlu pohledu bylo nutné mít širší znalosti o typech porovnávaných automobilů (například červená ŠKODA Fabia Combi v porovnání s červenou ŠKODA Fabia Hatchback), které nebyly na první pohled zřejmé. Ze získaných anotací si lze udělat dobrou představu o náročnosti vybrané datové sady. Zároveň je zde patrná užitečnost získaných ohodnocení, díky kterým bylo možné rozdělit datovou sadu podle sebejistoty oslovených anotátorů. Získané anotace byly také velkým přínosem pro stanovení



Obrázek 5.1: Webové rozhraní pomocí kterého byla shromažďována data zadaná uživateli za účelem vyhodnocení přesnosti algoritmu. Uživatelé klasifikovali zobrazený pár jako totožný/rozdílný. Jeden rozpoznávací test obsahoval 50 náhodně vybraných párů automobilů tak, aby celkové počty klasifikací na jeden pár byly rozloženy co nejrovnoměrněji.



Obrázek 5.2: Histogram zobrazující počet párů, jejichž průměrná klasifikace člověkem spadá do daného intervalu. Celkově bylo ohodnoceno 1 232 párů automobilů hodnotami 1 pro pozitivní (shodný) pár a -1 pro negativní (rozdílný) pár, z nichž každý byl ohodnocen přibližně 20×. Z grafu je patrné, že si lidé byli naprosto jisti ve 422 párech (222 negativních a 200 pozitivních).



Obrázek 5.3: ROC (receiver operating characteristic) zobrazující vztah mezi pravděpodobnostmi opravdu pozitivních (P(TP)) a falešně pozitivních (P(FP)) reidentifikací automobilů při použití natrénovaného lineárního klasifikátoru. V grafu jsou vykresleny 4 křivky pro různé prahy průměrných ohodnocení párů automobilů lidmi.

výchozích (ground truth) hodnot pro vyhodnocení úspěšnosti algoritmu. Vyhodnocení založené na principu anotací pomocí GUI aplikace uvedené v kapitole 4.4 by totiž jednoznačně nevypovídalo o přesnosti vytvořeného systému, jelikož z pohledu anonymní reidentifikace mohou být totožné i automobily, které nejsou identické. Takové případy jsou sice z pohledu reidentifikace nežádoucí, ale pro stanovení přesnosti vytvořeného reidentifikačního systému založeného pouze na vzhledu automobilů je nutné i tyto případy zahrnout.

Na základě získaných anotací bylo pak možné vyhodnotit přesnost natrénovaného lineárního klasifikátoru. Jeden ze způsobů vyhodnocení je pomocí ROC (receiver operating characteristic) [12] křivek, jak je znázorněno na obrázku 5.3. Křivka je tvořena z přesností klasifikátoru při nastavení určitého prahu regrese pro stejné/různé páry automobilů. Tato skutečnost je pak vyhodnocena mírou skutečně pozitivních reidentifikací (P(TP) - senzitivita) a mírou falešně pozitivních reidentifikací P(FP). Byly vypočítány čtyři křivky pro různé prahy pozitivních/negativních průměrných klasifikací, které byly získány od lidí. Modrá křivka byla například spočítána za použití intervalu $< -1, -0{,}6)$ pro negativní průměrnou lidskou klasifikaci a intervalu $(0{,}6, 1 >$ pro pozitivní průměrnou lidskou klasifikaci. Čím blíže je daný $\pm$ práh k nule, tím náročnější je datová sada pro lineární klasifikátor, jelikož obsahuje větší množství i pro člověka obtížných párů k reidentifikaci. Například je v grafu viditelný poměr 0,6/0,1 pro nulový práh (ohodnocená datová sada je rozdělena přesně v půlce v hodnotě 0 na stejné a nestejné páry), který vypovídá o tom, že algoritmus vyhodnotil 60 % ze všech pozitivních párů jako stejné a 10 % ze všech negativních párů jako

Obrázek 5.4: Ukázka párů algoritmem špatně vyhodnocena jako pozitivní (identická). Rozdíly ve vzhledu vozidel jsou velice malé (většinou v podobě jiného potisku na karoserii). I když jsou tyto rozdíly pro člověka evidentní a patrné, je velice náročné vytvořit lineární model pro regresi, který by je dokázal správně a hlavně v krátkém čase vyhodnotit.

falešně pozitivní. Nejlepších výsledků bylo dosaženo při stanovení váhy 0,4 pro výsledky regrese vycházející z barevných histogramů a váhy 1,6 pro HOG. Z výsledků dílčích regresí byl pak vybrán 90. percentil. Použitá datová sada vypovídá o pesimistickém odhadu přesnosti použitého klasifikátoru. Lze předpokládat, že v praxi by byla přesnost vyšší. Zejména pokud by systém vyhledával v menších databázích (v rámci stovek automobilů) na sebe navazujících a ne příliš od sebe vzdálených kamer. Přesnost výsledků viditelných v grafu svědčí o možnosti použití klasifikátoru pro statistické účely. Obrázek 5.4 ukazuje některé případy falešně pozitivních reidentifikací. V převážné většině případů se jedná o vzhledově velice podobná vozidla lišící se v malých (pro člověka k rozpoznání dostačujících) detailech.

Doba reidentifikace jednoho automobilu se odvíjí od velikosti množiny automobilů, ve kterých je hledaný automobil vyhledáván a na nastavení prahu regrese. Například při vyhledávání v databázi obsahující přibližně 400 vozidel trvá necelých 70 milisekund, než jsou pro daný automobil vypočítány všechny příznakové vektory. V průměru trvá pokus o nalezení korespondujícího vozidla dalších 15 milisekund. Uvedené hodnoty nejsou definitivní, ale pouze orientační. Při případném nasazení systému do každodenního používání jde zde pro konkrétní podmínky prostor pro optimalizaci algoritmu.

Zhodnocení a diskuse Na základě získaných poznatků mohu usoudit, že algoritmus dosahuje uspokojivých výsledků, které svědčí o použitelnosti systému v případném plném nasazení. I přes ne zcela zanedbatelnou chybovost je možné jeho využití například pro statistické účely v dopravě. Vytvořený systém není naprosto dokonalý a zůstává zde prostor pro mnohá vylepšení a rozšíření. Při podrobnější studii výsledků jsem zjistil, že se úspěšnost reidentifikací liší podle výskytu automobilů v jízdních pruzích. Reidentifikace automobilů jedoucích směrem ke kameře dosahovala menší chybovosti (falešně pozitivních výsledků) než u automobilů v opačném jízdním pruhu. Proto by bylo možné oddělit tyto dva typy reidentifikací nebo se zaměřit pouze na automobily směrující ke kameře. Dále by mohlo zvýšit přesnost systému sofistikovanější porovnávání jednotlivých stěn ohraničujícího kvádru, kdy výslednou regresi může značně ovlivnit špatná viditelnost některé z použitých stěn, přičemž jiné stěny jsou v obraze viditelné a ostré. Při nasazení systému do plného provozu by bylo vhodné kompaktněji uschovávat vypočítané příznaky pro rychlejší vyhledávání automobilů

a případnou efektivnější síťovou komunikaci v systému. Jak již bylo zmíněno dříve, díky anonymní reidentifikaci jsou některé případy pro lineární klasifikátor velice obtížně nebo téměř neřešitelné. Proto je pochopitelné, že přesnost systému úzce souvisí s unikátností vzhledových charakteristik snímaných automobilů. Systém selhává v případech, kdy se automobily liší velice málo (viz. obrázek 5.4). Naopak v sadě různorodých automobilů systém funguje poměrně spolehlivě.

Obrázek 5.5: Obrázek pokračuje obrázkem 5.6

Obrázek 5.6: Výběr z množiny 190 944 párů automobilů seřazených podle výsledků hluboké regrese (21 dílčích regresí, tj. podobnosti jak barevné tak tvarové) od největšího po nejmenší. Výsledky regrese jsou zobrazeny mezi jednotlivými páry. Čím větší číslo bylo regresí vypočítáno, tím více by se měl daný pár shodovat. Z obrázku je patrné, že se automobily začínají ve svém vzhledu shodovat přibližně od hodnoty 0,3 a výše. Velikost zobrazeného výběru byla vybrána tak, aby si čtenář mohl vytvořit dostatečnou představu o podobnosti různých párů automobilů a o výsledcích hluboké regrese k nim náležících.

Kapitola 6

Závěr

Tato práce se týkala problematiky reidentifikace automobilů v obraze vizuální cestou. Cílem práce bylo seznámení se s metodami a již existujícími přístupy (kapitoly 2 a 3), které tuto problematiku řeší. Na základě získaných poznatků pak navrhnout a experimentovat s algoritmy reidentifikace automobilů bez použití rozpoznání registrační značky a následně vybrané algoritmy implementovat do systému (kapitola 4). Vytvořený systém byl vyhodnocen na dostatečně velké datové sadě (kapitola 5).

Řešení práce spočívalo v extrakci maximálního množství potřebných informací o automobilu reprezentovaných jako příznakové vektory barevných RGB histogramů a histogramů orientovaných gradientů za použití 3D ohraničujícího kvádru pro detekci vozidla v obraze. Hlavní proces reidentifikace je založen na regresi vyhodnocené natrénovaným lineárním SVM klasifikátorem.

Systém byl vyhodnocen na 1232 podobných párech automobilů ohodnocených více jak 500 lidmi. Za použití uvedeného řešení bylo možné dosáhnout 60 % míry úspěšnosti správně pozitivních reidentifikací s 10 % mírou falešně pozitivních vyhodnocení na náročné datové sadě. Jedna reidentifikace v rozumně velké množině automobilů (v rámci stovek prvků) trvá přibližně 85 milisekund výpočetního výkonu procesoru (Core i7). Z toho vyplývá, že je možno plynule reidentifikovat přibližně 12 automobilů za sekundu.

Přesná reidentifikace automobilů ve videu přináší nový pohled na získávání dopravních informací, statistik a dalších důležitých dat v reálném čase. Některé otázky zůstaly ovšem nezodpovězeny a nechávají tak prostor pro zdokonalení systému. Uvedená kombinace příznaků dosahovala uspokojivých výsledků, které rozhodně stojí za pozornost, a může být hodnotnou inspirací pro podobně zaměřené práce. Řešení popsané v této práci má vysoký potenciál k tomu, aby bylo nasazeno do plného provozu v systémech sledující dopravní toky.

Práce byla prezentována na studentské konferenci Excel@FIT 2015, kde získala ohodnocení společnosti FEI. Dále byl o práci napsán vědecký článek na konferenci ITSC 2015. Rozhodnutí o přijetí nebylo doposud obdrženo.

Literatura

[1] Anagnostopoulos, C.; Alexandropoulos, T.; Loumos, V.; aj.: Intelligent traffic management through MPEG-7 vehicle flow surveillance. In *Modern Computing, 2006. JVA '06. IEEE John Vincent Atanasoff 2006 International Symposium on*, Oct 2006, s. 202–207, doi:10.1109/JVA.2006.30.

[2] Arth, C.; Leistner, C.; Bischof, H.: Object Reacquisition and Tracking in Large-Scale Smart Camera Networks. In *Distributed Smart Cameras, 2007. ICDSC '07. First ACM/IEEE International Conference on*, Sept 2007, s. 156–163, doi:10.1109/ICDSC.2007.4357519.

[3] Aziz, K.-E.; Merad, D.; Fertil, B.: People re-identification across multiple non-overlapping cameras system by appearance classification and silhouette part segmentation. In *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on*, Aug 2011, s. 303–308, doi:10.1109/AVSS.2011.6027341.

[4] Bao, S. Y.; Xiang, Y.; Savarese, S.: Object Co-detection. In *Proceedings of the 12th European Conference on Computer Vision - Volume Part I*, ECCV'12, Berlin, Heidelberg: Springer-Verlag, 2012, ISBN 978-3-642-33717-8, s. 86–101, doi:10.1007/978-3-642-33718-5˙7.

[5] Bradski, G.: *Learning OpenCV Computer Vision with the OpenCV Library*. Sebastopol: O'Reilly Media, Inc, 2008, ISBN 9780596554040.

[6] Charbonnier, S.; Pitton, A.; Vassilev, A.: Vehicle re-identification with a single magnetic sensor. In *Instrumentation and Measurement Technology Conference (I2MTC), 2012 IEEE International*, May 2012, ISSN 1091-5281, s. 380–385, doi:10.1109/I2MTC.2012.6229117.

[7] Chen, Z.; Chen, K.; Chen, J.: Vehicle and Pedestrian Detection Using Support Vector Machine and Histogram of Oriented Gradients Features. In *Computer Sciences and Applications (CSA), 2013 International Conference on*, Dec 2013, s. 365–368, doi:10.1109/CSA.2013.92.

[8] Dalal, N.; Triggs, B.: Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, ročník 1, June 2005, ISSN 1063-6919, s. 886–893 vol. 1, doi:10.1109/CVPR.2005.177.

[9] Dubská, M.; Herout, A.; Juránek, R.; aj.: Fully Automatic Roadside Camera Calibration for Traffic Surveillance. *IEEE Transactions on Intelligent Transportation Systems*, ročník PP, 2014, doi:10.1109/TITS.2014.2352854.

[10] Dubská, M.; Sochor, J.; Herout, A.: Automatic Camera Calibration for Traffic Understanding. In *Proceedings of the British Machine Vision Conference (BMVC)*, 2014.

[11] Fan, R.-E.; Chang, K.-W.; Hsieh, C.-J.; aj.: LIBLINEAR: A Library for Large Linear Classification. *Journal of Machine Learning Research*, ročník 9, 2008: s. 1871–1874.

[12] Fawcett, T.: An Introduction to ROC Analysis. *Pattern Recognition Letters*, ročník 27, č. 8, Červen 2006: s. 861–874, ISSN 0167-8655, doi:10.1016/j.patrec.2005.10.010.

[13] For, W.-K.; Leman, K.; Eng, H.-L.; aj.: A multi-camera collaboration framework for real-time vehicle detection and license plate recognition on highways. In *Intelligent Vehicles Symposium, 2008 IEEE*, June 2008, ISSN 1931-0587, s. 192–197, doi:10.1109/IVS.2008.4621176.

[14] Goto, Y.; Yamauchi, Y.; Fujiyoshi, H.: CS-HOG: Color similarity-based HOG. In *Frontiers of Computer Vision, (FCV), 2013 19th Korea-Japan Joint Workshop on*, Jan 2013, s. 266–271, doi:10.1109/FCV.2013.6485502.

[15] Hung, C.-C.; Peng, W.-C.: Model-Driven Traffic Data Acquisition in Vehicular Sensor Networks. In *Parallel Processing (ICPP), 2010 39th International Conference on*, Sept 2010, ISSN 0190-3918, s. 424–432, doi:10.1109/ICPP.2010.50.

[16] Kamijo, S.; Matsushita, Y.; Ikeuchi, K.; aj.: Traffic monitoring and accident detection at intersections. *Intelligent Transportation Systems, IEEE Transactions on*, ročník 1, č. 2, 2000: s. 108–118.

[17] Karatzoglou, A.; Meyer, D.; Hornik, K.: Support Vector Machines in R. *Journal of Statistical Software*, ročník 15, č. 9, 4 2006: s. 1–28, ISSN 1548-7660.

[18] Ke, Y.; Sukthankar, R.: PCA-SIFT: a more distinctive representation for local image descriptors. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, ročník 2, June 2004, ISSN 1063-6919, s. II–506–II–513 Vol.2, doi:10.1109/CVPR.2004.1315206.

[19] Lämmer, S.; Helbing, D.: Self-control of traffic lights and vehicle flows in urban road networks. *Journal of Statistical Mechanics: Theory and Experiment*, ročník 2008, č. 04, 2008, doi:10.1088/1742-5468/2008/04/P04019.

[20] Lee, H.; Kim, D.; Kim, D.; aj.: Real-time automatic vehicle management system using vehicle tracking and car plate number identification. In *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, ročník 2, July 2003, s. II–353–6 vol.2, doi:10.1109/ICME.2003.1221626.

[21] Naito, T.; Tsukada, T.; Yamada, K.; aj.: Robust license-plate recognition method for passing vehicles under outside environment. *Vehicular Technology, IEEE Transactions on*, ročník 49, č. 6, Nov 2000: s. 2309–2319, ISSN 0018-9545, doi:10.1109/25.901900.

[22] Ortuzar, J. d.; Willumsen, L. G.: *Modelling transport*. 2011.

[23] Press, W.: *Numerical recipes : the art of scientific computing*. Cambridge, UK New York: Cambridge University Press, 2007, ISBN 978-0-521-88068-8.

[24] Saghafi, M.; Hussain, A.; Saad, M.; aj.: Appearance-based methods in re-identification: A brief review. In *Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on*, March 2012, s. 404–408, doi:10.1109/CSPA.2012.6194758.

[25] Shapiro, L. G.; Stockman, G. C.: *Computer vision*. Upper Saddle River, NJ: Prentice Hall, 2001, ISBN 0-13-030796-3.

[26] Tan, T.; Kittler, J.: Colour texture analysis using colour histogram. *Vision, Image and Signal Processing, IEE Proceedings* -, ročník 141, č. 6, Dec 1994: s. 403–412, ISSN 1350-245X, doi:10.1049/ip-vis:19941420.

[27] Vallejo, D.; Albusac, J.; Jimenez, L.; aj.: A cognitive surveillance system for detecting incorrect traffic behaviors. *Expert Systems with Applications*, ročník 36, č. 7, 2009: s. 10503–10511.

[28] Vapnik, V.: *Statistical learning theory*. New York: Wiley, 1998, ISBN 978-0471030034.

[29] Wang, G.; Xiao, D.; Gu, J.: Review on vehicle detection based on video for traffic surveillance. In *Automation and Logistics, 2008. ICAL 2008. IEEE International Conference on*, Sept 2008, s. 2961–2966, doi:10.1109/ICAL.2008.4636684.

Příloha A

Plakát