

VĚDECKÉ SPISY VYSOKÉHO UČENÍ TECHNICKÉHO V BRNĚ

*Edice Habilitační a inaugurační spisy, sv. 498*

*ISSN 1213-418X*

**Adam Herout**

**KROKY SMĚREM  
K PLNĚ AUTOMATICKÉ  
VIZUÁLNÍ ANALÝZE DOPRAVY**

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta informačních technologií

Ústav počítačové grafiky a multimédií

**doc. Ing. Adam Herout, PhD.**

**KROKY SMĚREM K PLNĚ AUTOMATICKÉ  
VIZUÁLNÍ ANALÝZE DOPRAVY**

RECENT STEPS TOWARDS FULLY AUTOMATIC  
VISUAL TRAFFIC ANALYSIS

TEZE PŘEDNÁŠKY  
K PROFESORSKÉMU JMENOVACÍMU ŘÍZENÍ  
V OBORU  
VÝPOČETNÍ TECHNIKA A INFORMATIKA



BRNO 2015

## **KLÍČOVÁ SLOVA**

Počítačové vidění, Monitorování dopravy, Automatická kalibrace kamery, Houghova transformace, Detekce přímk, Detekce úběžníků, Manhattanské světy, Měření rychlosti, Měření vzdálenosti, Analýza dopravy.

## **KEYWORDS**

Computer Vision, Traffic Surveillance, Automatic Camera Calibration, Hough Transform, Line Detection, Detection of Vanishing Points, Manhattan Worlds, Speed Measurement, Distance Measurement, Traffic Analysis.

© Adam Herout, 2015

ISBN 978-80-214-5135-3

ISSN 1213-418X

# Obsah

<b>Představení autora</b>	<b>4</b>
<b>1 Úvod</b>	<b>5</b>
<b>2 Detekce přímek pomocí duality přímka <math>\leftrightarrow</math> bod</b>	<b>6</b>
2.1 Houghova transformace pro detekci přímek v obraze . . . . .	6
2.2 Paralelní souřadnice – jemný úvod . . . . .	8
2.3 Detekce přímek pomocí parametrizace PClines . . . . .	9
<b>3 Detekce úběžníků v „diamantovém prostoru“</b>	<b>11</b>
3.1 Kaskádová aplikace transformace PClines . . . . .	11
3.2 Diamantový prostor . . . . .	14
<b>4 Plně automatická kalibrace kamery v dopravě</b>	<b>15</b>
4.1 Kalibrace kamery – určení úběžníků . . . . .	15
4.2 Konstrukce trojrozměrných obalových kvádrů . . . . .	16
4.3 Statistická adaptace rozměrů vozidel . . . . .	17
<b>5 Extrakce obrazových vzorků automobilů</b>	<b>20</b>
5.1 Akvizice vzorků vozidel . . . . .	21
5.2 Extrakce příznaků pro shlukování vzorků vozidel . . . . .	22
5.3 Cenová funkce rozdílnosti vozidel . . . . .	23
5.4 Automatické shlukování vzorků aut . . . . .	24
5.5 Vzájemná registrace shluků vozidel . . . . .	24
5.6 Předběžné výsledky . . . . .	25
<b>6 Závěr</b>	<b>27</b>
<b>Literatura</b>	<b>28</b>

## Představení autora



Po absolvování inženýrského studia na Fakultě informačních technologií VUT v Brně doc. Adam Herout pokračoval ve studiu v doktorském studijním programu Výpočetní technika a informatika, kde v roce 2004 obhájil disertační práci s názvem „Hardware Architecture for Point-Based Graphics Rendering“. Od roku 2001 je zaměstnancem FIT VUT v Brně. Nejdříve v rámci doktorského studia, po dokončení doktorského studia v roce 2004 jako odborný asistent a po obhájení habilitační práce „Hardware Acceleration of Graphics and Vision Algorithms“ v roce 2010 působí jako docent. Od roku 2007 působí jako zástupce vedoucího Ústavu počítačové grafiky a multimédií.

Výsledky své vědecké a odborné práce publikoval doc. Herout ve 14 časopiseckých publikacích s IF, v 50 sbornících světových konferencí, i jinde. Vypracoval desítky posudků zahraničních publikací, domácích publikací nebo projektů a posudků disertačních prací. Publikované práce byly 358× citovány, z toho je 36 citací v databázi Web of Science. Za FIT je hlavním řešitelem 2 zahraničních projektů a podílel se na 3 dalších v roli spoluřešitele. Byl nebo je řešitelem nebo spoluřešitelem 7 domácích projektů.

Doc. Herout je uznávaným vysokoškolským pedagogem. Vyučoval řadu předmětů, zavedl 2 předměty v magisterském studijním programu. Organizuje četné extrakurikulární aktivity pro studenty, kteří je oceňují a nad rámec svých studijních povinností je navštěvují. Na svém blogu píše rady pro studenty, zejména řešitele diplomových prací a jeho texty využívá velké množství studentů z různých vysokých škol v ČR a SR. V současné době je školitelem 10 studentů doktorského studijního programu Výpočetní technika a informatika. Dva studenti doktorského studijního programu pod jeho vedením již úspěšně obhájili disertační práci, z toho jedna práce byla oceněna stipendiem pro nejlepší disertační práci v akademickém roce 2014/2015. Byl vedoucím téměř 200 bakalářských a diplomových prací z nichž některé byly oceněny prvními místy v soutěžích ACM SPY, Diplomka roku, Student EEICT, a četné získaly další ocenění a byly publikovány na konferencích.

# 1 Úvod

Nelze pochybovat o přínosu počítačů, který stojí na jejich schopnosti informace uchovávat, vyhledávat a zpřístupňovat uživatelům. Aby měly co uchovávat, vyhledávat a zpřístupňovat, potřebují počítače informace i vhodným způsobem získávat – tato jejich schopnost je podmínkou nutnou jejich užitečnosti. Mnoho informací, s nimiž pracují, získávají počítače z reálného světa pomocí uživatelských rozhraní – to jest lidský uživatel poskytuje počítači informace předem navrženou formou tak, aby jim stroj přímo rozuměl. Užitečnost počítačů a komunikačních sítí se ovšem dramaticky zvyšuje se schopností strojů informace o reálném světě získávat přímo pomocí sensorů – počítačovým viděním, analýzou zvuku a z dalších modalit.

Tento text popisuje některé autorovy výzkumné práce přinášející pokroky v počítačovém vidění, tj. získávání hodnotných informací z obrazu nebo videa. Konkrétně je tento text zaměřen na získávání informací pozorováním automobilové dopravy dohledovými kamerami. Značným omezením existujících systémů řešících tento problém je, že lidský operátor musí před použitím jednotlivou dohledovou kameru zkalibrovat, tj. dodat vhodné informace o scéně a o kameře, které si algoritmus nedokáže zjistit sám. Aplikování dohledového systému na zvolenou kameru není tedy okamžité, ale je podmíněno jistým množstvím kvalifikované lidské práce a v některých případech i provedení reálných měření ve scéně zabírané kamerou. Podobná operace je potřebná i při re-aplikaci systému na kameru, u které došlo ke změně pozice nebo nastavení její optiky. Práce shrnuté v tomto textu směřují k automatizaci kalibrace a přinášejí nové poznatky klíčové pro vytvoření plně automatického systému, který nevyžaduje vůbec žádný lidský vstup.

Text nejprve popisuje matematickou parametrizaci přímk PCLines založenou na paralelních souřadnicích. Tato parametrizace umožňuje detekovat přímky pomocí Houghovy transformace inovovaným způsobem tak, že jsou umožněny další operace, které mnohé jiné parametrizace přímk neumožňují. Těchto specifických vlastností využívá kaskádová aplikace transformace PCLines popisovaná dále v textu. Toto kaskádové použití řečené transformace umožňuje efektivně detekovat v obraze úběžníky. Na detekci úběžníků v textu navazuje návrh jejich specifického použití – pro charakterizaci proudu přímočaře se pohybujících vozidel v dopravě. V textu (a dříve publikovaných pracích) ukazujeme, že proud dopravy je možné charakterizovat trojicí navzájem ortogonálních úběžníků, že pozorováním dohledového videa dopravy je možné úběžníky přesně a spolehlivě určit a nakonec, že trojice ortogonálních úběžníků takto získaných definuje kalibraci kamery. Dalším krokem je určení měřítka scény, jež není součástí vnitřních ani vnějších parametrů kamery, ale je nezbytné pro měření vzdáleností a rychlostí v jednotkách reálného světa (typicky v metrech či v kilometrech za hodinu). Text je završen popisem probíhajícího a dosud nepublikovaného výzkumu, kde zmiňované navržené metody jsou využity k vytvoření datové sady pro učení přesné klasifikace automobilů. Na rozdíl od existujících přístupů je datová sada tvořena opět převážně bez použití lidského vstupu. Použití popisovaných technik vede k tvorbě trénovacích a testovacích dat bezprecedentních rozměrů, s minimálním lidským úsilím.

Autor využívá této příležitosti k vyjádření vděčnosti svým spolupracovníkům – studentům doktorského studia, jež měl a má tu čest a potěšení vést a již jsou spoluautory děl shrnutých v tomto textu. Bez jejich talentů a úsilí by práce popisované v tomto textu nemohly vzniknout. Díky, Markéto, Jirko a Kubo!

## 2 Detekce přímek pomocí duality přímka $\leftrightarrow$ bod

Houghova transformace [1] je známý a dlouho studovaný nástroj pro detekci tvarů a objektů v rastrových obrazech. Hough ji původně použil k detekci přímek, ale později byla zobecněna pro komplexnější tvary – kružnice, elipsy a dokonce obecné vzory [2]. Princen et al. [3] Houghovu transformaci formalizoval a popsal jako proces testování hypotéz.

### 2.1 Houghova transformace pro detekci přímek v obraze

Houghova transformace je typicky používána k detekci křivek s analytickým popisem. V takovém případě jsou jako *důkazy* přítomnosti objektu používány hrany detekované ve vstupním rastrovém obraze vhodným operátorem (Sobelův, Prewittův, ...). *Hypotézy* o hledaném objektu jsou možné křivky dané třídy popsané příslušným počtem parametrů. Například přímka má dva, kružnice tři stupně volnosti v 2D prostoru.

Rodina (2D) křivek může být specifikována implicitní funkcí

$$f(x, y, p_1, \dots, p_N) = f(\mathbf{x}, \mathbf{p}) = 0, \quad (2.1)$$

kde  $x$  a  $y$  jsou souřadnice v obrázku a  $p_1, \dots, p_N$  jsou souřadnice v *parametrickém prostoru*. Houghova transformace převádí bod  $(x, y)$  na takovou část parametrického prostoru (křivku), která popisuje všechny možné křivky dané třídy procházející daným bodem. Souřadnice v parametrickém prostoru a jejich zobrazení na křivku určují *parametrizaci* dané třídy křivek (jedna třída může mít velký počet různých parametrizací). Algoritmus 1 pak popisuje detekci křivky specifikované rovnicí (2.1) –  $\Omega$  je množina všech pixelů v obraze,  $P_H$  je množina všech souřadnic v parametrickém prostoru  $H$ .

---

**Algoritmus 1** Detekce implicitních křivek pomocí Houghovy transformace

---

**Vstup:** Vstupní obraz  $I$ , rozměry parametrického prostoru  $H$

**Výstup:** Detekované křivky  $C$

$H(\mathbf{x}) \leftarrow 0, \forall \mathbf{x} \in P_H$

**for all**  $\mathbf{x} \in \Omega$  **do**

**if**  $\mathbf{x}$  je hrana v  $I$  **then**

**for all**  $\{\mathbf{p} : \mathbf{p} \in P_H, f(\mathbf{x}, \mathbf{p}) = 0\}$  **do**

$H(\mathbf{p}) \leftarrow H(\mathbf{p}) + 1$

**end for**

**end if**

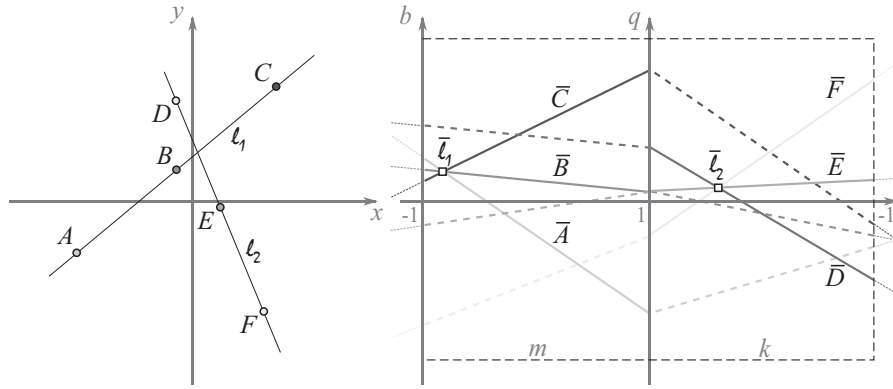
**end for**

$C = \{\mathbf{p} | \mathbf{p} \in P_H, \text{na pozici } \mathbf{p} \text{ je vysoké lokální maximum v } H\}$

---

Houghova transformace vznikla jako nástroj pro detekci přímek a je v tomto kontextu stále velmi populární. Existuje velký počet parametrizací přímek a algoritmických variací. 2D přímky jsou polynomy řádu 1 a mají tedy 2 stupně volnosti; všechny parametrizace jsou tudíž dvourozměrné. Obecně, nechť je přímka reprezentována svým normálovým vektorem  $\mathbf{n} = (n_x, n_y)$  a vzdáleností od počátku souřadnic  $\varrho$ . Každý bod  $\mathbf{p} = (p_x, p_y)$  ležící na přímce musí splňovat rovnici

$$\mathbf{p} \cdot \mathbf{n} = \varrho. \quad (2.2)$$



Obrázek 2.1: Houghova transformace založená na  $m$ - $b$  a  $k$ - $q$  parameterizacích přímky (2.5). **vlevo:** prostor vstupního obrázku; **vpravo:** odpovídající Houghův prostor.

Hough v roce 1962 [1] vyvinul transformaci s parametrizací přímky ve formě *slope-intercept* (byť sama parametrizace v původním patentu výslovně uvedena není). Tato forma se obvykle zapisuje:

$$\ell : y = xm + b, \quad (2.3)$$

nicméně metoda popisovaná v Houghově patentu odpovídá parametrizaci (obrázek 2.1):

$$\ell : x = ym + b. \quad (2.4)$$

Při použití této parametrizace všechny přímky procházející jedním bodem utvářejí přímku v parametrickém prostoru, tudíž se jedná o zobrazení bod-na-přímku (*point-to-line mapping, PTLM*). Jako v případě kteréhokoliv PTLM je parametrický prostor všech možných čar existujících v omezeném vstupním obrázku nekonečný [4]. Pro získání konečného parametrického prostoru je potřeba použít alespoň dvou komplementárních prostorů parametrů. V případě  $m - b$  parametrizace (2.3) mohou být tyto dvě dílčí parametrizace například:

$$\begin{aligned} y &= xm + b, \\ x &= yk + q. \end{aligned} \quad (2.5)$$

Tuytelaars et al. [5] přidali třetí komplementární prostor a použili vzniklou trojitou parametrizaci pro detekci úběžníků a horizontu. Tato modifikace jménem *Cascaded Hough Transform* používá tři páry parametrů založené na rovnici

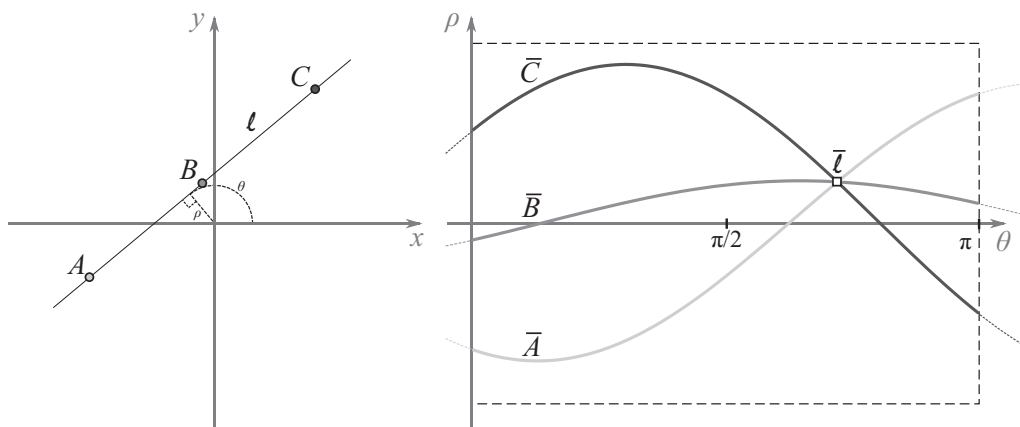
$$ax + b + y = 0. \quad (2.6)$$

První z nich má souřadnice  $a$ - $b$ , druhý  $(1/a)$ - $(b/a)$  a třetí  $(1/b)$ - $(a/b)$ . Všechny prostory jsou omezeny do intervalu  $< -1, 1 >$  v obou směrech. I pro neomezený vstupní obraz je pro každou přímku definován bod v jednom z takto omezených tří prostorů. Více informací o CHT je k nalezení u autorů [5, 6].

V roce 1972 Duda a Hart [7] představili velice populární parametrizaci označovanou  $\theta$ - $\rho$ , jež je velice důležitá proto, že její parametrický prostor je přirozeně konečný, bez nutnosti konstrukce vícero komplementárních podprostorů. Je založená na rovnici přímky v normální formě:

$$y \sin \theta + x \cos \theta = \rho, \quad (2.7)$$





Obrázek 2.2: Houghova transformace používající  $\theta$ - $\rho$  parametrizaci přímky. **vlevo:** vstupní omezený prostor; **vpravo:** odpovídající Houghův prostor.

kde parametr  $\theta$  je úhel náklonu přímky a  $\rho$  je vzdálenost přímky od počátku souřadného systému (obrázek 2.2 vlevo). Při použití této parametrizace obrazy všech přímek procházejících daným bodem leží na sinusoidě v parametrickém prostoru (obrázek 2.2 vpravo).

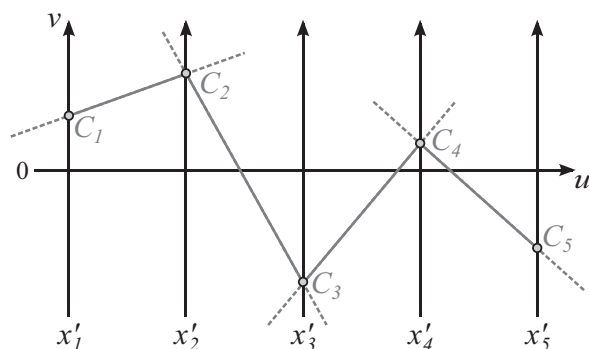
Podobně jako parametrizace  $\theta$ - $\rho$  vychází *Circle Transform* [8] z rovnice přímky v normální formě (2.7). Místo parametrů  $\theta$  a  $\rho$  ovšem používá jako parametry souřadnice průsečíku dané přímky a její normály procházející počátkem souřadnic. Další konečná parametrizace je *Muff transform* představená Wallacem v roce 1985 [9]. Jako základ této parametrizace slouží obdélník ohraničující vstupní obrázek. Každá přímka procházející obrázkem se s tímto obdélníkem protíná přesně ve dvou bodech. Vzdálenost prvního z nich podél obvodu obdélníku od počátku je prvním parametrem, druhým pak vzájemná vzdálenost prvního a druhého průsečíku, opět podél obvodu obdélníku. Použití obalové kružnice místo obdélníku definuje další omezenou parametrizaci *fan-beam* [10]. Každá relevantní přímka protíná tuto kružnici opět ve dvou bodech. Úhly definované těmito body byly jako parametry Houghovy transformace použity poprvé Eckhardtem and Maderlechnerem [11]. Transformace *Muff transform* je základem parametrizace navržené Formanem [12], jenž ji zkombinoval s parametrizací  $\theta$ - $\rho$  a reprezentoval přímky prvním průsečíkem s obalovým obdélníkem a jejich sklonem  $\theta$ .

## 2.2 Paralelní souřadnice – jemný úvod

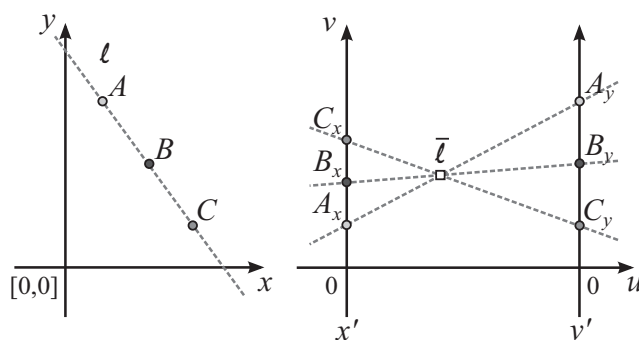
Obvyklým způsobem vizualizace vektorů v  $N$ -rozměrném euklidovském prostoru je použit kartézské souřadnice, kde každý vektor je reprezentován přesně jedním bodem v  $N$ -rozměrném souřadném systému. Ovšem vizualizace prostorů s více než dvěma dimenzemi na obvyklém dvourozměrném prostředku kreslení (papír, obrazovka počítače, ...) je možná pouze projekcí, kde dochází ke snížení intuitivnosti a množství prezentovaných informací.

Systém paralelních souřadnic [13] reprezentuje vektorový prostor pomocí os, jež jsou vzájemně rovnoběžné. Každý  $N$ -rozměrný vektor je pak reprezentován  $N - 1$  přímkami spojující tyto osy – obrázek 2.3. V tomto textu použijeme euklidovskou rovinu s kartézským souřadným systémem  $u$ - $v$  pro definování pozic bodů v prostoru paralelních souřadnic. Pro určení těchto bodů použijeme notaci  $(u, v, w)_{\mathbb{P}^2}$  využívající homogenních souřadnic v projektivním prostoru  $\mathbb{P}^2$  a  $(u, v)_{\mathbb{E}^2}$  užíjeme pro kartézské souřadnice v euklidovském prostoru  $\mathbb{E}^2$ .

Ve dvourozměrném případě jsou body v prostoru  $x$ - $y$  reprezentovány jako přímky v prostoru paralelních souřadnic. Reprezentace kolineárních bodů (ležících na jedné přímce) se protínají v jednom bodě – reprezentaci příslušné přímky (obrázek 2.4). Na základě této relace



Obrázek 2.3: Reprezentace 5-rozměrného vektoru v paralelních souřadnicích. Vektor je reprezentován svými souřadnicemi  $C_1, \dots, C_5$  na osách  $x'_1, \dots, x'_5$ , spojenými úplnou polyčarou (složenou ze 4 nekonečných přímek).



Obrázek 2.4: Tři kolineární body v paralelních souřadnicích: **vlevo:** prostor kartézských souřadnic, **vpravo:** prostor paralelních souřadnic. Přímka  $\ell$  je v paralelních souřadnicích reprezentována bodem  $\bar{\ell}$ .

je možné definovat zobrazení přímka  $\leftrightarrow$  bod (PTLM, point-to-line-mapping) mezi těmito prostory. V některých případech, jako u přímky  $\ell : y = x$ , leží odpovídající bod  $\bar{\ell}$  v nekonečnu (je to ideální bod). Projektivní prostor  $\mathbb{P}^2$  (na rozdíl od euklidovského prostoru  $\mathbb{E}^2$ ) poskytuje vyjádření souřadnicemi i pro tyto zvláštní případy. Vztah mezi přímkou  $\ell : ax + by + c = 0$  (zapisovanou  $[a, b, c]$ ) a bodem, jež ji reprezentuje  $\bar{\ell}$ , je definován zobrazením:

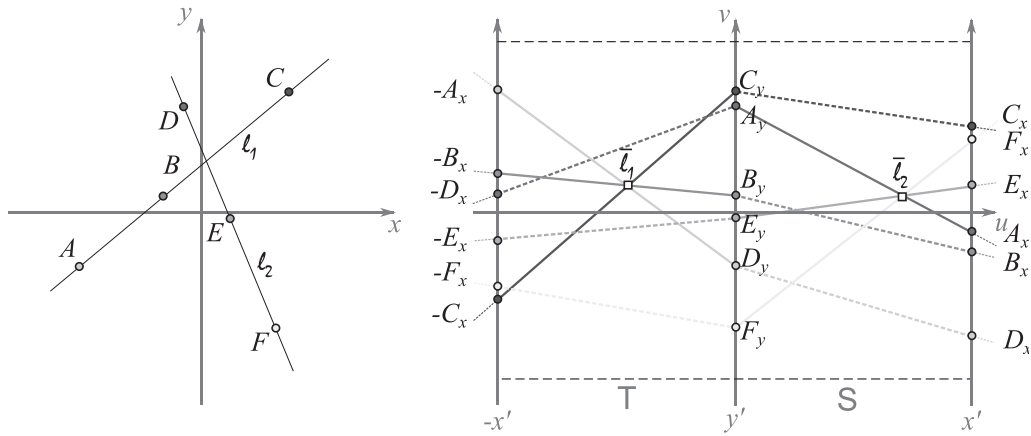
$$\ell : [a, b, c] \rightarrow \bar{\ell} : (db, -c, a + b)_{\mathbb{P}^2}, \quad (2.8)$$

kde  $d$  je vzdálenost mezi rovnoběžnými osami  $x'$  a  $y'$ .

## 2.3 Detekce přímek pomocí parametrizace PClines

Díky tomu, že paralelní souřadnice můžou sloužit jako zobrazení přímka  $\leftrightarrow$  bod (PTLM), mohou být použity jako parametrizace přímek pro Houghovu transformaci. Parametrizaci PClines založenou na této myšlence jsme vyvinuli a publikovali s kolegy [14]. Velice podobný postup byl zvolen autory El Mejdani et al. [15].

Přímka  $\ell : y = mx + b$  je v paralelních souřadnicích (prostor  $u-v$ ) reprezentována bodem  $\bar{\ell} = (d, b, 1 - m)_{\mathbb{P}^2}$ . Tento bod leží mezi osami  $x'$  a  $y'$  právě tehdy, když  $-\infty < m < 0$ . Pro  $m = 1$  je  $\bar{\ell}$  ideálním bodem (bodem v nekonečnu). Pro  $m = 0$  leží  $\bar{\ell}$  na ose  $y'$ , v případě svislých přímek ( $m = \pm\infty$ ) leží  $\bar{\ell}$  na ose  $x'$ . Kromě tohoto prostoru paralelních souřadnic  $x', y'$  (označovaného jako *straight*,  $\mathcal{S}$ , přímý), jsme navrhli použít komplementární prostor *twisted* ( $\mathcal{T}$ , zkroucený)  $x', -y'$ , jenž je identický, kromě toho, že osa  $y$  je invertována. V tomto komplementárním prostoru leží  $\bar{\ell}$  mezi osami  $x'$  a  $-y'$  právě tehdy, když  $0 < m < \infty$ . Zkombinováním



Obrázek 2.5: **vlevo:** Původní prostor  $x-y$ , **vpravo:** odpovídající  $\mathcal{TS}$  prostor.

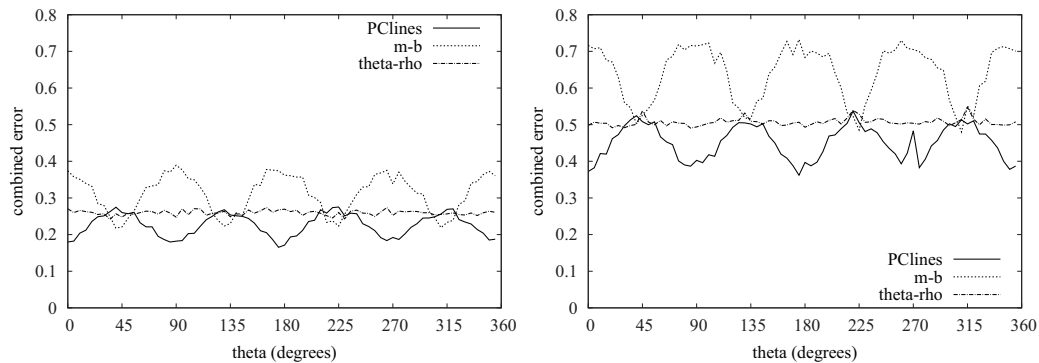
obou prostorů vznikne celá rovina  $\mathcal{TS}$  – obrázek 2.5. Použitím obou prostorů postačuje k reprezentaci konečného vstupního obrázku končená část roviny  $u-v$ :

$$\begin{aligned}
 & -d \leq u \leq d, \\
 & -\max\left(\frac{W}{2}, \frac{H}{2}\right) \leq v \leq \max\left(\frac{W}{2}, \frac{H}{2}\right),
 \end{aligned}
 \tag{2.9}$$

kde  $W$  a  $H$  jsou šířka a výška vstupního rastrového obrázku.

Pro samotnou detekci přímek je použit standardní algoritmus Houghovy transformace: prostor  $u-v$  omezený podle (2.9) je uniformně vzorkován a tvoří tak matici akumulátorů; vstupní obrázek je zpracován; pro všechny pixely (nebo vhodně omezenou podmnožinu pixelů) s hledanou vlastností, například s vysokou odezvou hranového detektoru, je příslušná množina akumulátorů inkrementována. V případě parametrizace PClines je do duálního prostoru vždy rasterizována dvojice úseček – do každé z polovin prostoru  $\mathcal{TS}$ . Vodorovné souřadnice koncových bodů jsou fixní  $\{0, d, -d\}$ ; svislé souřadnice jsou přímo souřadnice  $x, y$  a  $-y$  zpracovávaného bodu.

Obrázek 2.6 srovnává chyby tří důležitých parametrizací ( $m-b$ ,  $\theta-\rho$ , PClines) pro přímky v závislosti na jejich sklonu  $\theta$ . Pro každý  $5^\circ$  interval sklonu bylo generováno 100 syntetických obrázků, každý s jednou přímkou a 20 000 šumovými pixely. Vyhodnocena byla průměrná chyba pro všechny obrázky (obr. 2.6 vlevo) a průměrná chyba pro 5 nejmeně přesně detekovaných čar (tj. pesimistický odhad, obr. 2.6 vpravo).



Obrázek 2.6: Chyba lokalizace přímek v závislosti na sklonu přímky. Sklony jsou rozděleny do skupin o šířce  $5^\circ$ . **vlevo:** průměrná chyba přes všechny přímky; **vpravo:** průměrná chyba přes 5 nejhorších pokusů.

### 3 Detekce úběžníků v „diamantovém prostoru“

Úběžníky jsou důležitou charakteristikou mnoha reálných scén. V *manhattanských světech* mohou být použity pro určení orientace kamery [16, 17], pro její kalibraci [18], rekonstrukci scény [19], pro shlukování hran ve scéně za účelem porozumění scéně [16] a pro další úlohy [20].

Detekce úběžníků v obraze má za sebou dlouhou historii. Populárním nástrojem vždy byla Houghova transformace [21]; nejprve za použití Gaussovy koule [22], následně v různých modifikacích. Tuytelaars et al. [23] vyvinuli *Cascaded Hough Transform* (CHT) – zobecněnou lineární Houghovu transformaci. Li et al. [24] použili 1D Houghovu transformaci kaskádovanou jiným způsobem pro oddělenou detekci úběžníků přítomných v obraze a úběžníků vzdálených.

Bazin et al. [25] upozornili, že přístupy založené na Houghově transformaci detekují úběžníky odděleně a nedokážou detekovat (optimální) trojice ortogonálních úběžníků, které jsou užitečné pro interpretaci snímků manhattanských světů. Bazin et al. navrhuje hledat v okolí iniciálních hypotéz o trojicích ortogonálních úběžníků mechanickým prohledáváním nebo pomocí algoritmu RANSAC [26]. RANSAC a jeho modifikace jsou v detekci úběžníků populární [27, 28, 26, 29]. Některé práce se vyrovnávají s úběžníky v nekonečnu (ideálními body) tím, že zvlášť hledají úběžníky konečné a ideální [30]. Další používaná rodina algoritmů je založená na *expectation maximization* (EM) [31, 16].

V naší práci [32] navazující na parametrizaci PClines (kapitola 2) jsme navrhli algoritmus pro detekci ortogonálních trojic úběžníků. Navržený přístup je podobný CHT [23], ale v celém procesu je používán jediný akumulací prostor a akumulace se děje pouze jednou. Podobně jako u PClines, parametrizace přímk (a úběžníků) je lineární – algoritmus nevyžaduje použití goniometrických nebo dalších nelineárních operací.

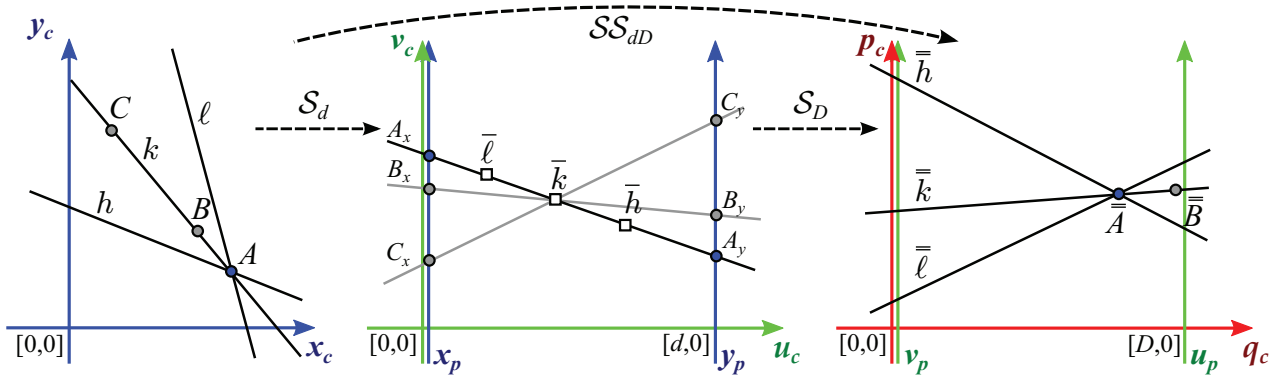
#### 3.1 Kaskádová aplikace transformace PClines

Náš algoritmus kaskádové aplikace parametrizace PClines vychází z důkazu o kaskádovatelnosti PClines podobně jako v CHT [33]. Hlavní myšlenka CHT (*Cascaded Hough Transform*) [23] je nejdříve akumulovat data o hranách v obraze do duálního obrazu, tam detekovat maxima (odpovídající přímkám) a tato maxima akumulovat opět do dalšího Houghova prostoru. Díky tomu, že obrazy přímk protínajících se v jednom bodě leží na přímce, v druhém Houghově prostoru se jejich obrazy protínají v jednom bodě – obraze úběžníku.

V následujícím textu bude index  $p$  označovat souřadné osy v prostoru paralelních souřadnic a index  $c$  osy kartézského souřadného systému. Pro odlišení bodů a přímk bude reprezentace  $[x, y, w]$  používána pro 2D bod v homogenních souřadnicích a  $(a, b, c)$  označí přímku definovanou rovnicí  $ax + by + c = 0$ . Jak bylo odvozeno v předchozí kapitole, zobrazení parametrizací PClines (viz obrázek 3.1 uprostřed) je určeno následovně:

$$\begin{aligned} \mathcal{S}_d^p([x, y, w]) &= (-x + y, -dw, dx), \\ \mathcal{S}_d^l((a, b, c)) &= [db, -c, a + b]. \end{aligned} \tag{3.1}$$

V duchu CHT může být Houghova transformace použita znovu pro nalezení úběžníků, tj. body jsou znovu zobrazeny na přímky a přímky na body. Nechť  $D$  označuje vzdálenost paralelních souřadných os v druhém zobrazení. Složené zobrazení vzniklé dvojí aplikací rovnice (3.1) je pak:



Obrázek 3.1: Kaskádová aplikace dvou transformací PCLines s použitím *přímého* prostoru  $\mathcal{S}$ . **vlevo:** Prostor vstupního obrázku s body a přímkami. **uprostřed:** Stejné objekty v prostoru paralelních souřadnic. Zde je definován nový kartézský souřadný systém (zeleně,  $u_c, v_c$ ). **vpravo:** Druhá transformace do prostoru paralelních souřadnic  $u_p, v_p$ .

$$\mathcal{SS}_{dD}^p([x, y, w]) = (\mathcal{S}_D^l \circ \mathcal{S}_d^p)([x, y, w]) = [-dDw, -dx, -x + y - dw], \quad (3.2)$$

$$\mathcal{SS}_{dD}^l((a, b, c)) = (\mathcal{S}_D^p \circ \mathcal{S}_d^l)((a, b, c)) = (db + c, Da + Db, -dDb). \quad (3.3)$$

Dvojice zobrazení  $\mathcal{SS}_{dD}$  jsou transformacemi jednoho nekonečného prostoru do druhého nekonečného prostoru [4]. V případě použití pro detekci přímek v konečném obraze je možné cílový nekonečný prostor nahradit dvojicí komplementárních konečných prostorů (kapitola 2, [14, 23]). Mapování do komplementárního *zkrouceného* prostoru  $\mathcal{T}$  je:

$$\begin{aligned} \mathcal{T}_d^p([x, y, w]) &= (x + y, -dw, dx), \\ \mathcal{T}_d^l((a, b, c)) &= [db, -c, a - b]. \end{aligned} \quad (3.4)$$

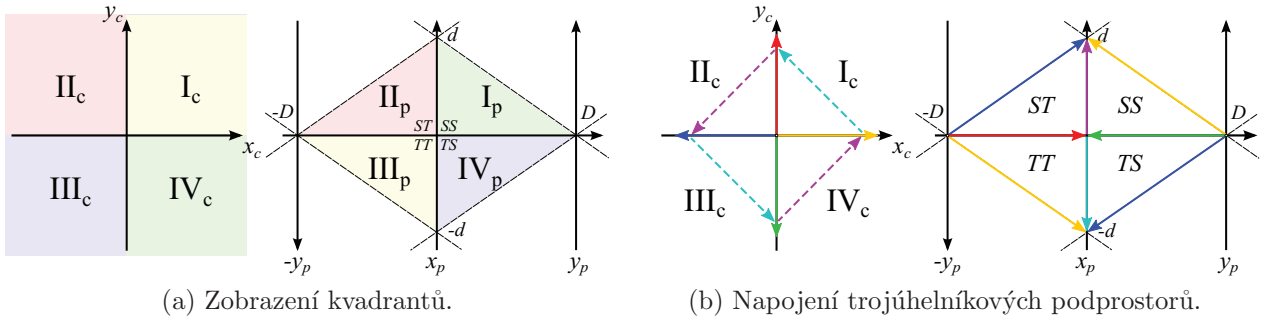
Přímý a zkroucený prostor je v kaskádově aplikovaných zobrazeních možno aplikovat všemi čtyřmi možnými způsoby:

$$\begin{aligned} \mathcal{SS}_{dD}^p([x, y, w]) &= (\mathcal{S}_D^l \circ \mathcal{S}_d^p)([x, y, w]) = [-dDw, -dx, -x + y - dw], \\ \mathcal{ST}_{dD}^p([x, y, w]) &= (\mathcal{S}_D^l \circ \mathcal{T}_d^p)([x, y, w]) = [-dDw, -dx, -x + y + dw], \\ \mathcal{TS}_{dD}^p([x, y, w]) &= (\mathcal{T}_D^l \circ \mathcal{S}_d^p)([x, y, w]) = [-dDw, -dx, x + y - dw], \\ \mathcal{TT}_{dD}^p([x, y, w]) &= (\mathcal{T}_D^l \circ \mathcal{T}_d^p)([x, y, w]) = [-dDw, -dx, x + y + dw]. \end{aligned} \quad (3.5)$$

Použití všech čtyř kombinací transformuje celou (vstupní) projektivní rovinu do konečného prostoru složeného ze čtyř konečných částí (každá pro jednu transformaci) – viz obrázek 3.2a.

Například je možno uvážit transformaci (3.2) a body z kvadrantu  $IV_c$  (obr. 3.2a). Konečný bod, tj.  $P = [x, y, 1]$ ,  $x \geq 0$ ,  $y \leq 0$ , je transformován na  $\mathcal{SS}_{dD}^p(P) = [-dD, -dx, -x + y - d]$ . Protože  $d, D > 0$  (vzdálenosti mezi paralelními souřadnými osami), pozice obrazu daného bodu může být odvozena takto:

$$\begin{aligned} d, D > 0, x \geq 0 &\Rightarrow -dD < 0, -dx \leq 0 \\ x \geq 0, y \leq 0, d > 0 &\Rightarrow -x + y - d < 0 \\ &\Rightarrow \frac{-dD}{-x + y - d} > 0 \\ &\Rightarrow \frac{-dx}{-x + y - d} \geq 0. \end{aligned} \quad (3.6)$$



Obrázek 3.2: Podprostory zobrazení. **(a) vlevo:** Kvadranty původního nekonečného kartézského prostoru. **vpravo:** Kvadranty prostoru PClines (dva spojené prostory paralelních souřadnic). **(b) vlevo:** Po kaskádové aplikaci dvou zobrazení jsou cílové kvadranty omezené a mají tvar trojúhelníku. **vpravo:** Tyto trojúhelníkové plochy mohou být přeuspořádány a napojeny na hranách, jež reprezentují identické body v původním prostoru. Čárkované úsečky zastupují ideální body v původním prostoru; dohromady reprezentují jednu ideální přímku celé projektivní roviny (prostor  $x_c, y_c$  je nekonečný).

Tím je dokázáno, že při použití transformace  $\mathcal{SS}$  jsou všechny body z kvadrantu  $IV_c$  transformovány do kvadrantu  $I_p$ . Navíc obrazy všech těchto bodů leží v jedné polorovině definované přímkou  $\ell_b = (d, D, -dD)$  (přímka procházející body  $[D, 0]$  a  $[0, d]$ ):

$$(d, D, -dD) \cdot \left[ \frac{-dD}{-x + y - d}, \frac{-dx}{-x + y - d}, 1 \right] = \frac{-dDy}{-x + y - d} \leq 0. \quad (3.7)$$

Obrazy  $\mathcal{SS}_{dD}^p(P) = [0, -dx, -x + y]$  všech ideálních bodů  $P = [x, y, 0]$ ,  $x \geq 0$ ,  $y \leq 0$  leží na úsečce spojující body  $[0, 0]$  a  $[0, d]$ .

Tudíž *všechny* body – konečné i ideální – z kvadrantu  $IV_c$  jsou transformovány zobrazením  $\mathcal{SS}$  do průniku kvadrantu  $I_p$  a uzavřenou dolní polorovinou definovanou přímkou  $dx + Dy - dD = 0$ . Důkazy pro zobrazení (3.5) jsou analogické.

Tyto čtyři části, každá zobrazená jiným zobrazením ( $\mathcal{SS}, \mathcal{ST}, \mathcal{TS}$ , nebo  $\mathcal{TT}$ ), mohou být navzájem vhodně napojeny, protože obrazy os  $\mathbf{x}_c, \mathbf{y}_c$  a ideální přímka vždy leží na spojených stranách jednotlivých segmentů (obrázek 3.2b). Při takovémto uspořádání je zobrazení bodů z původní roviny do spojeného *diamantového prostoru*:

$$[x, y, w]_o \rightarrow [-dDw, -dx, \text{sgn}(xy)x + y + \text{sgn}(y)dw]_d \quad (3.8)$$

$$[x, y, w]_d \rightarrow [Dy, \text{sgn}(x)dx + \text{sgn}(y)Dy - dDw, x]_o, \quad (3.9)$$

kde  $\text{sgn}$  je nenulová funkce signum. Výsledkem transformace přímky je polyčára, jejíž počet segmentů závisí na počtu kvadrantů, kterými přímka prochází. Koncové body této polyčáry jsou určeny rovnicí (3.10). V případě, že přímka prochází pouze dvěma kvadranty (svislé či vodorovné přímky, přímky skrze počátek souřadnic), jeden segment po transformaci vždy degeneruje do bodu (tj. dva následující body podle rovnice (3.10) jsou totožné). Singulární případy, tj. přímky  $x = 0$ ,  $y = 0$  a ideální přímka, po transformaci vystupují, jak je ukázáno na obrázku 3.2b.

$$\alpha = \text{sgn}(ab), \quad \beta = \text{sgn}(bc), \quad \gamma = \text{sgn}(ac) \quad (3.10)$$

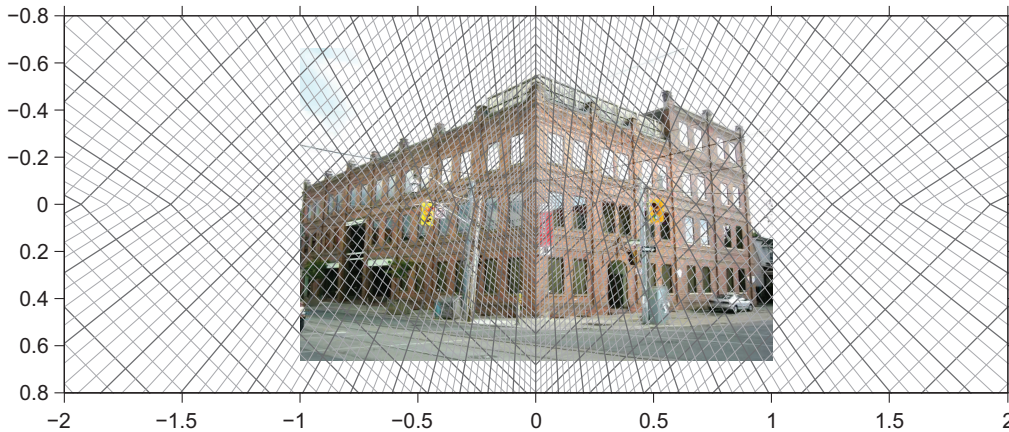
$$(a, b, c)_o \rightarrow \left[ \frac{\alpha a}{c + \gamma a}, \frac{-\alpha c}{c + \gamma a} \right], \left[ \frac{b}{c + \beta b}, 0 \right], \left[ 0, \frac{b}{a + \alpha b} \right], \left[ \frac{-\alpha a}{c + \gamma a}, \frac{\alpha c}{c + \gamma a} \right]$$

## 3.2 Diamantový prostor

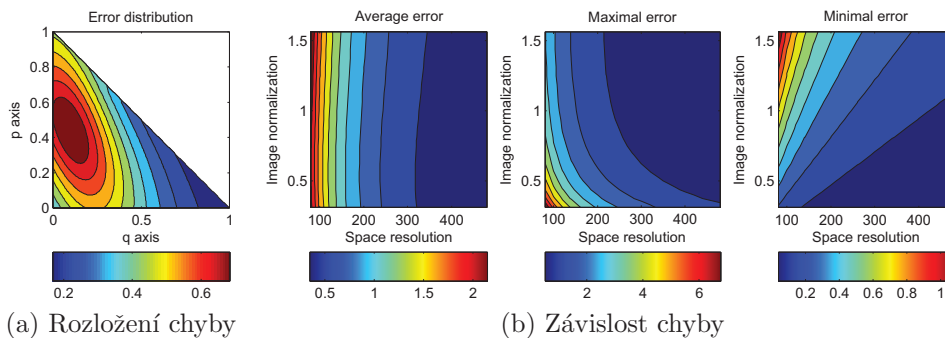
Cílový prostor zobrazení (obr. 3.2) má tvar „diamantu“ a proto ho v tomto textu a jiných publikacích označujeme jako *diamantový prostor* (diamond space). V konkrétních implementacích je možné trojúhelníkové podprostory přeuspořádat pro dosažení úspornější organizace paměti.

Obrázek 3.3 ukazuje diskretizaci prostoru akumulátorů, když je představené zobrazení používáno v duchu Houghovy transformace pro detekci úběžníků. Tvar zpětně transformovaných akumulátorů či „pixelů“ je jiný ve směru osy  $x$  a  $y$ , protože diamantový prostor byl vytvořen invertováním osy  $y$  – toto je dáno libovůlí tvůrce a při jiné volbě by tvar akumulátorů byl mírně odlišný.

Zvolením jiné normalizace vstupního obrázku a/nebo rozlišení prostoru akumulátorů je možné nalézt vhodné parametry pro dosažení přijatelné diskretizační chyby. Diskretizace prostoru akumulátorů je dána rozlišením prostoru, normalizací vstupního obrázku a vnitřními parametry kamery. Smysluplnou metrikou pro hodnocení diskretizační chyby pro libovolnou buňku akumulátoru je maximum odchylek všech vektorů promítnutých do dané buňky od reprezentativního vektoru vypočteného pro danou buňku. Tato chyba je vykreslena v obrázku 3.4.



Obrázek 3.3: Vizualizace akumulátorů („pixelů“) promítnutá z diamantového prostoru zpět do roviny vstupního obrázku. Normalizace obrázku  $\mu = 1$ .



Obrázek 3.4: Chyba v prostoru akumulátorů:  $d = D = 1$ , parametry kamery a obrázku jsou získány z York Urban DB [16]. (a) Distribuce úhlové chyby pro jednu čtvrtinu diamantového prostoru s rozlišením  $320^2$  a normalizací  $\mu = 1$ . (b) Závislost průměrné/maximální/minimální chyby v diamantovém prostoru na jeho rozlišení a na normalizaci obrázu. Rozlišení diamantového prostoru je kritické pro průměrnou chybu a změnou normalizace je možné ohraničit nejvyšší a nejnižší chybu.

## 4 Plně automatická kalibrace kamery v dopravě

Automatický kamerový dohled je užitečný v organizaci dopravy – pro sběr statistických dat [34], pro přímé řízení dopravy [35], pro vynucení práva [36, 37], atd. Existující systémy obvykle vyžadují manuální nastavení lidským operátorem, často vyžadující fyzická měření v zájmové scéně [38]. Naším cílem je zpracovávat dopravní data plně automaticky, bez jakéhokoli lidského vstupu. Toto zahrnuje zjištění vnitřních parametrů kamery, vnějších parametrů ve vztahu k proudu dopravy a měřítko roviny po které se doprava pohybuje, což umožní měřit v jednotkách reálného světa – obrázek 4.6.

Některé existující práce v této oblasti vyžadují expertní vstup ve formě anotace vodorovného dopravního značení se známou šířkou pruhů [39] nebo známými rozměry značení [40], definování pozice kamery [39, 41], průměrné velikosti vozidla [42, 43] nebo průměrné rychlosti vozidel [44]. Společným rysem prakticky všech metod je detekce úběžníku korespondujícího se směrem proudu dopravy (plná kalibrace kamery by vyžadovala tři ortogonální úběžníky [19, 45, 46]). Populárním přístupem je získávat tento úběžník z čar na silnici [47, 48, 49] nebo z jízdnic pruhů [47, 50, 51], více nebo méně automaticky získaných z obrazu. Tyto přístupy obvykle vyžadují vysoký počet jízdnic pruhů a konzistentní a dobře viditelné vodorovné dopravní značení. Schoepflin a Dailey [44] pozorováním vozidel segmentují jednotlivé pruhy; opět je zapotřebí velký počet pruhů – dálnice apod. Jiní výzkumníci detekují vozidla detektorem založeným na boostingu a sledují jejich pohyb [52], nebo analyzují hrany přítomné na vozidlech [53]. Beymer et al. [54] akumulují sledované význačné body a zároveň vyžadují uživatelské zadání vhodných čar. Kanhere a Birchfield [55] použili zvláštní přístup pro kamery umístěné nízko nad úrovní vozovky. Když je jednou provedena kalibrace kamery a je známo měřítko scény, může být rektifikována rovina vozovky, což umožní nejrůznější aplikace [56, 57, 41, 40, 58].

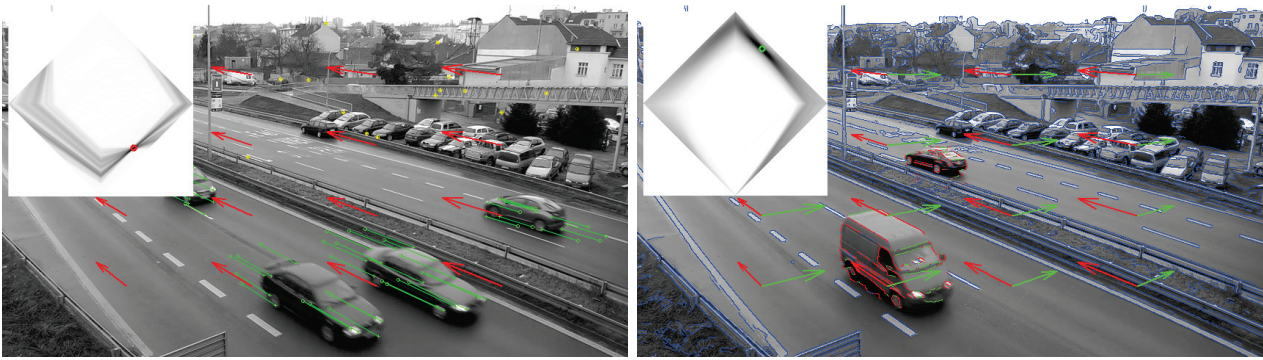
Podobně jako ostatní autoři, v našem přístupu předpokládáme středové promítání (*pinhole camera*). Na rozdíl od ostatních přístupů nepředpokládáme vodorovný horizont [44, 56, 50]. Toto omezení by bylo příliš limitující a naše práce ho nevyžaduje, protože jsme navrhli metodu pro přesné nalezení druhého úběžníku definujícího scénu (kapitola 4.1). Při této kalibraci předpokládáme nulové radiální zkreslení kamery, ale v naší předchozí práci [59] je problém radiálního zkreslení vyřešen.

### 4.1 Kalibrace kamery – určení úběžníků

Naše metoda kalibrace kamery používaná pro monitorování dopravy umožňuje získat ohniskovou vzdálenost a orientaci kamery vzhledem k proudu dopravy. Je založena na detekci dvou vzájemně ortogonálních směrů – *první* je směr pohybu vozidel a *druhý* je na něj kolmý a rovnoběžný s rovinou vozovky. Za předpokladu, že optický střed (principal point) je ve středu projekční roviny, je možné dopočítat *třetí* ortogonální směr a získat ohniskovou vzdálenost kamery.

Pro detekci prvního úběžníku jsou v obraze detekovány výrazné body a tyto jsou sledovány pomocí KLT trackeru v následujících snímcích. Ty z nich, které vykazují dostatečný pohyb, jsou považovány za možné úsečky ve směru pohybu vozidel a akumulovány do diamantového prostoru. Akumulátor s největším počtem hlasů je považován za zobrazení prvního úběžníku (obrázek 4.1 vlevo).





Obrázek 4.1: **vlevo:** Sledované body použité k určení prvního úběžníku. Body vyznačené zeleně se dostatečně pohybují a proto jsou akumulovány. Body vyznačené žlutě jsou stabilní a neúčastní se volby. Akumulační diamantový prostor je vlevo nahoře. **vpravo:** Akumulace druhého úběžníku. Modré hrany patří do pozadí, červené hrany jsou vyloučeny z akumulace protože jsou svislé nebo směřují k prvnímu úběžníku. Zelené hrany jsou akumulovány do diamantového prostoru (vlevo nahoře, zelený kroužek označuje maximum).

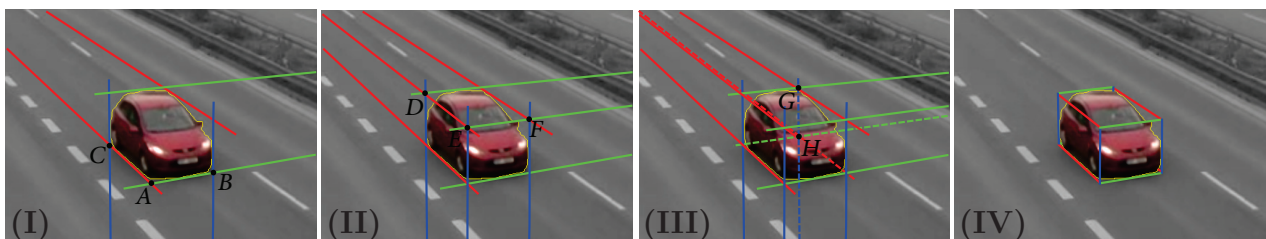
Druhý úběžník je také nalezen akumulací do diamantového prostoru – na základě pozorování, že mnoho hran na projíždějících vozidlech směřuje ve směru tohoto úběžníku. K detekci hran na vozidlech je použit hranový model pozadí. V něm je pro každý pixel uložena pravděpodobnost výskytu orientované hrany (orientace hran jsou diskretizovány do osmi směrů). Z hran popředí jsou odfiltrovány ty, které směřují k prvnímu úběžníku (známému z předchozího kroku) a svislé hrany. Hrany, které projdou těmito testy jsou akumulovány pro nalezení druhého úběžníku (obrázek 4.1 vpravo).

## 4.2 Konstrukce trojrozměrných obalových kvádrů

Dalším krokem je konstrukce trojrozměrných obalových kvádrů pozorovaných vozidel (jako na obrázku 4.2 (IV)). Zde zpracováváme siluety vozidel určených pomocí modelování pozadí a detekce popředí [60, 61]. Tvary vozidel je možné určit spolehlivě, včetně odstranění stínů [62]. Dále předpokládáme, že většina zájmových vozidel se pohybuje směrem k/od prvního úběžníku. Pohyblivé objekty, které tento předpoklad nespĺňují, vylučujeme z dalšího zpracování.

Vycházíme z pozorování, že některé části obrysu vozidel jsou velmi stabilní a spolehlivé – obrázek 4.2. Červené přímky procházejí prvním úběžníkem a jsou tečné k obrysu vozidla. Zelené tečny procházejí druhým úběžníkem a modré třetím.

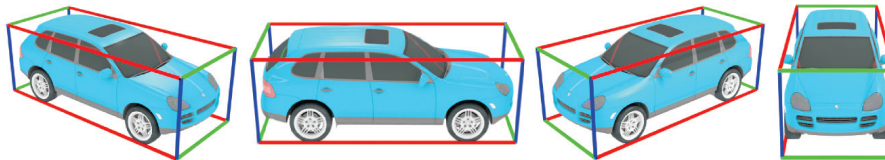
Protože obrysy vozidel nejsou zcela přesné a vozidla nejsou přesné kvádry, konstrukce oba-



Obrázek 4.2: Konstrukce trojrozměrného obalového kvádru vozidel. (I) Tečné přímky a jejich relevantní průsečíky  $A, B, C$ . (II) Odvozené přímky a jejich průsečíky  $E, D, F$ . (III) Odvozené přímky a průsečík  $H$ . (IV) Vytvořený obalový kvádr.

lových kvádrů je nejednoznačná, tj. záleží na pořadí, v němž jsou konstruovány tečny a jejich průsečíky. Navrhujeme následující pořadí, které se jeví jako nejstabilnější. Nejprve je zkonstruován bod  $A$  jako průsečík spodní červené a zelené tečny. Body  $B$  a  $C$  jsou definovány jako průsečíky spodní zelené a pravé modré tečny, resp. spodní červené a levé modré, obrázek 4.2 (I). Zkonstruované úsečky  $AB$  a  $AC$  definují kratší a delší stranu základny obalového kvádrů. Bod  $D$  je průsečíkem horní zelené tečny a levé modré tečny. Společně s přímkou procházející bodem  $A$  a třetím úběžníkem jednoznačně definuje bod  $E$ , obrázek 4.2 (II). Bod  $E$  může být také zkonstruován pomocí bodu  $F$ , což by vedlo k alternativní pozici bodu  $E$ . Volíme bod  $E$  s větší vzdáleností  $|AE|$ , což zajistí, že celá silueta vozidla bude zahrnuta v obalovém kvádrů. Se znalostí bodů  $F$  a  $D$  je možné zkonstruovat bod  $G$  jako průsečík přímky procházející bodem  $D$  a druhým úběžníkem s přímkou procházející bodem  $F$  a prvním úběžníkem, obrázek 4.2 (III).

Při odlišných konfiguracích úběžníků vzhledem k těžišti detekovaného automobilu se množina a pořadí použitých tečen a konstruovaných bodů drobně odlišuje. Úprava konstrukce je zřejmá a následuje stejné principy jako popsaná konstrukce. Obrázek 4.3 ukazuje ostatní možné orientace obalového kvádrů vzhledem k různým konfiguracím úběžníků.

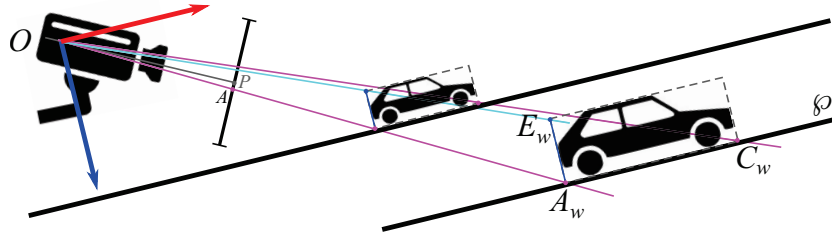


Obrázek 4.3: Různá uspořádání obalových kvádrů v závislosti na pozicích úběžníků vzhledem ke kameře. Kvůli zaoblení obrysu auta hrany obalového kvádrů někdy nepřiléhají těsně k detekované siluete vozidla. V praxi lze ovšem pozorovat, že nejméně jeden směr kvádrů přiléhá k siluete vozidla dostatečně přesně, což je dostačující pro určení měřítka scény.

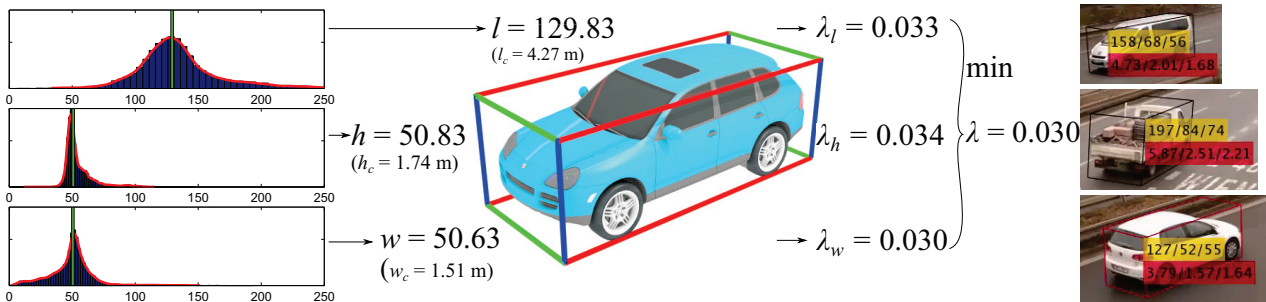
Střecha a obrys vozidla jsou obvykle poněkud zaoblené, takže detekovaný obalový kvádr je o trochu menší než odpovídá skutečným rozměrům vozidla. S touto nepřesností počítáme ve fázi určení měřítka (kapitola 4.3) a využíváme nejpřesněji přiléhající dvojici protilehlých stěn obalového kvádrů. Experimenty ukazují [63], že výsledná přesnost není tímto jevem ovlivněna. Pro určení správných obalových kvádrů je třeba odstranit z tvaru detekovaného auta jeho případný stín. Robustní řešení odstranění stínů přesahuje rámec této práce, ale je popisováno jinými výzkumníky [64, 65]. V naší práci předpokládáme přítomnost pouze měkkých stínů a pro jejich odstranění používáme metodu Horpraserta et al. [62].

### 4.3 Statistická adaptace rozměrů vozidel

Při znalosti promítnutého obalového kvádrů vozidla je přímo možné spočítat trojrozměrné souřadnice a rozměry tohoto kvádrů, až na přesné měřítko (obrázek 4.4). Uvažujeme trojrozměrný kartézský souřadný systém s kamerou umístěnou do  $O = [p_x, p_y, 0]$ , středem promítací roviny  $P = [p_x, p_y, f]$  (kde  $[p_x, p_y]$  je optický střed) a tři ortogonální směry odvozené z úběžníků detekovaných v obraze. Nejprve je zkonstruována rovina  $\varphi$  rovnoběžná s rovinou vozovky – její orientace je známa, protože směr třetího úběžníku je na ni kolmý. Vzdálenost této roviny od kamery je zvolena libovolně. Obrázek 4.4 ukazuje dvě možná umístění této roviny a vliv jejího umístění – čím blíže je rovina kamery, tím menší jsou zkonstruované objekty (scény jsou vzájemně geometricky podobné, přičemž správném měřítko vůči světovým jednotkám je dosud neznámé). Detekované vrcholy obalového kvádrů (body  $A, B, C, E$ ) jsou



Obrázek 4.4: Výpočet světových souřadnic. Rovina  $\varphi$  je rovnoběžná s rovinou vozovky a je odvozená na základě detekovaných úběžníků. Její vzdálenost od kamery je zvolena libovolně a přesné měřítko je odvozeno později, viz obrázek 4.5. Kamera je umístěna do  $O = [p_x, p_y, 0]$  a body základny obalového kvádrů (ve světových souřadnicích) jsou průsečíky roviny  $\varphi$  s paprsky z  $O$  skrze body  $A, C$  (zkonstruované dříve v projekční rovině). Ostatní body jsou průsečíky paprsků z  $O$  skrze promítnuté body a paprsky kolmé k  $\varphi$  procházející body  $A_w, B_w, C_w, H_w$ .



Obrázek 4.5: Výpočet měřítka scény  $\lambda$ . **vlevo:** Pro každý z rozměrů ( $l, w, h$ ) je určen jeho medián (zelený sloupec). **uprostřed:** Měřítka jsou určena zvlášť pro jednotlivé rozměry ze známých mediánů rozměrů vozidel ( $l_c, w_c, h_c$ ) jako  $\lambda_l = l_c/l$ ;  $\lambda_w = w_c/w$ ;  $\lambda_h = h_c/h$ . Výsledné měřítko je určeno jako minimum těchto tří hodnot. **vpravo:** Příklady relativních rozměrů vozidel (žlutě) a reálné rozměry v metrech po nalezení měřítka  $\lambda$  (červeně).

na rovinu  $\varphi$  promítnuty:

$$\begin{aligned} A_w &= \varphi \cap \overleftrightarrow{OA}, & B_w &= \varphi \cap \overleftrightarrow{OB}, & C_w &= \varphi \cap \overleftrightarrow{OC}, \\ E_w &= p_E \cap \overleftrightarrow{OE}; & p_E &\perp \varphi \wedge A_w \in p_E. \end{aligned} \quad (4.1)$$

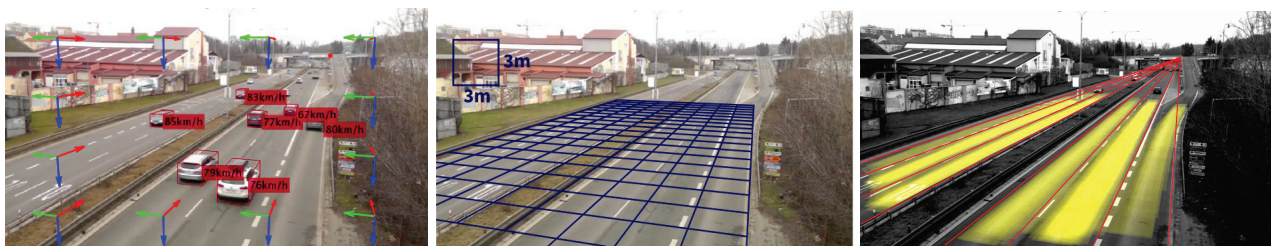
Na základě znalosti světových souřadnic obalového kvádrů je možné určit jeho rozměry (v neznámém, ale pro danou kameru a scénu konstantním měřítku):

$$(l, w, h) = (|A_w C_w|, |A_w B_w|, |A_w E_w|). \quad (4.2)$$

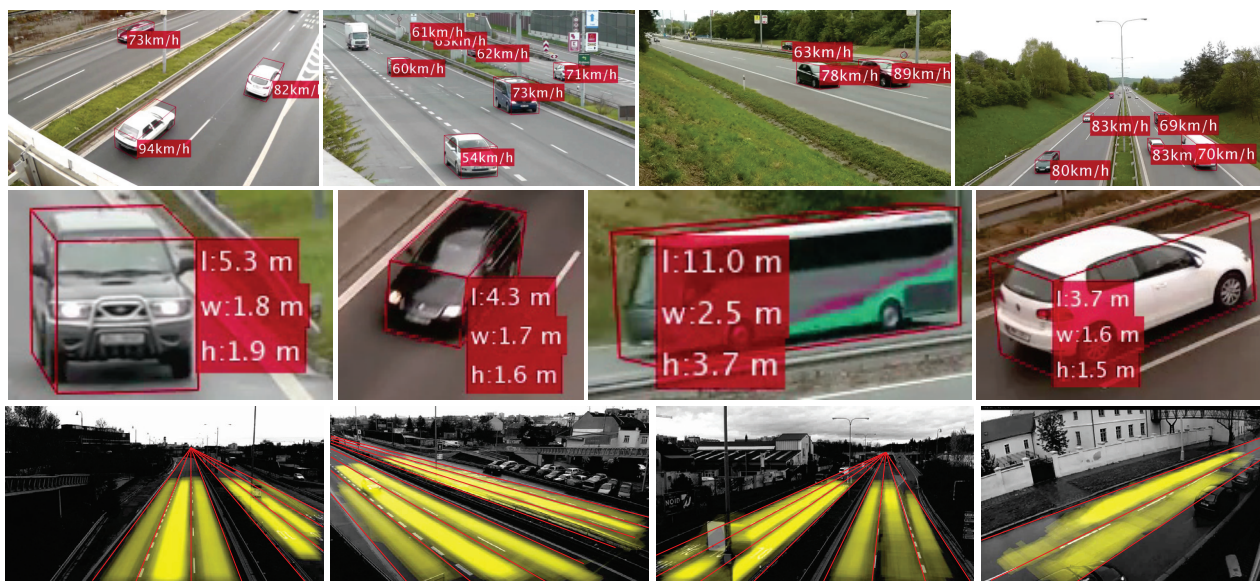
Zbývá nalézt měřítko  $\lambda$ , aby bylo lze určit skutečné rozměry ve světových jednotkách  $(\mathbf{l}, \mathbf{w}, \mathbf{h}) = \lambda(l, w, h)$ . Za tímto účelem jsme shromáždili statistická data o prodávaných autech a jejich rozměrech a vytvořili jsme histogram rozměrů jejich obalových kvádrů. Relativní rozměry vozidel ( $l, w, h$ ) jsou rovněž akumulovány do odpovídajícího histogramu. Histogramy potvrzují intuitivní předpoklad, že vozidla mají poměrně jednotnou výšku a šířku (vrcholek histogramu je výraznější), ale různí se ve své délce. Nalezením shody statistiky známých rozměrů s měřenými daty je možné získat hledané měřítko  $\lambda$ , obrázek 4.5.

Obrázek 4.6 demonstruje možné aplikace popisovaného algoritmu. Obrázek vlevo zobrazuje detekované úběžníky a detekovaná vozidla s vykreslenými obalovými kvádry a se spočtenou okamžitou rychlostí jejich pohybu. Obrázek uprostřed znázorňuje kalibraci v rovině vozovky

pomocí sítě čtverců zvolených rozměrů ( $3 \times 3$  metry) promítnutých z roviny vozovky do pohledu kamery. Kalibraci v rovině vozovky je možné použít pro měření rychlosti, měření rozměrů vozidel a pro další aplikace. Obrázek vpravo ukazuje výsledek akumulace pozorovaných podstav pohyblivých automobilů pro určení jízdnic pruhů a jejich dominantních směrů pohybu. Na obrázku 4.7 pak je několik příkladů scén s určenými rychlostmi a rozměry vozidel. Přesnost měření rychlosti je cca 3%, jak ukázala série měření v reálných dopravních scénách [63].



Obrázek 4.6: Prezentovaný algoritmus určuje 3 ortogonální úběžníky a konstruuje obalové kvádry (*vlevo*), načež následně automaticky určí měřítko promítání pomocí známých statistik rozměrů vozidel. Na základě těchto informací je možné měřit rozměry a rychlost (*uprostřed*) a analyzovat dopravní scénu (*vpravo*).



Obrázek 4.7: Příklady dosažených výsledků. **nahore:** Různé scény s automaticky měřenou rychlostí pohybu. **uprostřed:** Výřezy vozidel s určenými rozměry. **dole:** Jízdní pruhy detekované pomocí určených trojrozměrných obalových kvádrů.

## 5 Extrakce obrazových vzorků automobilů

V poslední době získává na pozornosti problém přesného rozpoznání výrobce a modelu vozidel. Většina přístupů se omezuje na jeden směr pohledu a nebo úzké spektrum pohledů [66, 67, 68, 69, 70]. Až velice nedávno, Hsiao et al. [71] navrhli rozpoznávač modelu automobilu pracující z víceřech pohledů. Svůj rozpoznávač učí na synteticky zobrazených datech (stále omezených na pohledy z pozorovatele stojícího v rovině vozovky) osmi různých modelů výrobce Honda. Podobná řešení jsou stále omezená trénovacími a testovacími daty – velikosti dostupných datových sad jsou spíše skromné.

Sběr datových sad reálných obrázků pro přesnou klasifikaci vozidel je obtížným úkolem. Stark et al. [72] shromáždili datovou sadu o 14 kategoriích (celkově 1904 vzorků) a komentují: *“We note that the data set is heavily biased w.r.t. viewpoints, which reflects the availability of images we encountered during data collection. It proved almost impossible to collect more than a handful of images for certain combinations of car-type and viewpoint.”* Krause et al. [73] zmiňují další problémy při sběru větší datové sady (16 185 obrázků, 197 tříd). Prohledávají populární webové stránky o autech a analyzují modely vozidel. Na základě seznamu modelů vozidel získali obrázky s použitím služeb Flickr, Google a Bing a anotovali/ověřili datovou sadu ručně. Mezi problémy, jimž čelili, uvádějí: *“The class list of cars changes on a yearly basis, the appearance of some models of cars remains constant from year to year, and typical car websites may even list cars which differ only in terms of non-visual features, e.g. their engine, as separate classes.”* Směry pohledu na vozidla jsou také rozloženy velice nerovnoměrně – označené obrázky vozidel umístěné online fanoušky aut jsou obvykle obrázky pořízené z vybraných a obvyklých pohledů. Lin et al. [74] shromáždili svoji datovou sadu 300 obrázků 30 modelů vozidel, protože v předcházejících datových sadách postrádali anotaci význačných bodů (landmarks) na vzorcích aut.

Naším cílem je přistoupit novým způsobem k problému shromáždění velké a přesné datové sady vozidel pokrývající všechny směry pohledu relevantní pro dohled v dopravě. Z dlouhodobého sledování dopravy dohledovými kamerami získáváme vizuální třídy zcela automaticky, bez lidského vstupu, viz obrázek 5.1. Klíčovým bodem zde je rozlišit mezi velice



Obrázek 5.1: Vizuální shlukování vozidel bez učitele ze dvou dohledových kamer. Každý řádek reprezentuje jeden shluk automaticky získaný z 10 000 pozorování vozidel získaných během 50 dní sledování jedné silnice.

jemnými vizuálními rozdíly a zároveň zahrnout do jednoho vzorku vozidla za různého osvětlení, v různých prostředích, atd.

Mnoho existujících přístupů k přesnému rozpoznání typu vozidla je omezených na frontální nebo zadní pohled a jsou založeny na detekci registrační značky a extrakci zájmové oblasti kolem ní [66, 67, 68, 69]. Autoři těchto prací používají různá schémata pro extrakci příznakových vektorů a pro samu klasifikaci. Stark et al. [72] používají přesnou klasifikaci modelu vozidla pro získání metrických informací a pro přibližné určení hloubky ve scéně z jednotlivých obrázků (obsahujících vozidla v použitelných pózách). Krause et al. [75] rekonstruují trojrozměrné modely vozidel pomocí stereo rekonstrukce se širokou základnou a kombinují rozpoznané trojrozměrné modely s dvourozměrnými reprezentacemi (Locality Constrained Linear Coding, LLC [76] a BubbleBank [77]), což jim umožní provádět přesnou klasifikaci vozidel.

Specifický přístup zvolili Thakoor a Bhanu [78], když počítají strukturní signatury vozidel odpovídající jejich geometrickému tvaru. Jejich vstupem je signál z jednotlivých dopravních dohledových kamer (zadní pohled na vozidla) a v tomto videu klasifikují projíždějící vozidla na základě jejich tvaru. Další přístup, navržený Prokajem a Medionim [79], je založen na určení pózy vozidla a je schopen zpracovat takřka libovolný pohled. Autoři navrhují použít trojrozměrné modely vozidel, uzpůsobit je rozpoznané póze, promítnout do roviny pohledu a použít příznaky založené na SIFT pro srovnání vozidel.

V naší práci navrhujeme shromáždit přesnou datovou sadu vozidel zcela bez ručního vstupu (bez učitele). Zpracováváme běžící video ze statických dohledových kamer a sledujeme projíždějící vozidla. Pro každé sledované vozidlo získáme několik (3–4) reprezentativních obrázků na předdefinovaných pozicích. Tyto skupiny obrázků jsou analyzovány po odpovídajících dvojicích pomocí cenové funkce jejich vizuální rozdílnosti (založené na SED [80] a na ICP [81]) a jsou shlukovány do shluků společného vzhledu. Shlukování je prováděno konzervativním způsobem, takže pouze vzorky s velice malou rozdílností padnou do jednoho shluku. Při jiném nastavení ovšem může být shlukování daleko tolerantnější a formovat rozsáhlé shluky pouze přibližně podobných vozidel. Rozdělení do shluků s různou úrovní tolerance rozdílnosti umožní formovat stromovou strukturu, jež může být anotována lidským operátorem pro získání asociací shluků s konkrétními označeními výrobce a modelu vozidla. Tento proces vyžaduje podstatně méně lidské práce v porovnání s anotováním jednotlivých obrázků vozidel (což je existující standard) a vede k potenciálně výrazně větším počtům anotovaných vzorků. Společně se získáním těchto anotací systém automaticky ke vzorkům připojuje informaci o směru pohledu a měřítku, která je získávána pro každý vzorek při sběru. Jak bylo vysvětleno v předchozí kapitole 4, směr pohledu a velikost vozidla můžou být z dohledových kamer získány zcela automaticky.

## 5.1 Akvizice vzorků vozidel

Obrázky (vzorky) vozidel pro shlukování jsou získány automaticky z dohledových kamer v dopravě. Kolem detekovaných aut jsou konstruovány trojrozměrné obalové kvádry (kapitola 4.2). Za vzorky pro shlukování jsou vybrány takové pozice vozidel, kde je střed podstavy vozidla nejbližší ručně vytipovaným pozicím v záběru kamery (viz obrázek 5.2). Množina takto definovaných míst v obraze bude označována  $\mathbb{P}$  a jednotlivá taková pozice  $\alpha \in \mathbb{P}$ . Pro každý ze vzorků jsou také získány a zaznamenány rozměry a maska vozidla získaná zpracováním vstupního videa [82, 83]. Dále jsou zpracována pouze vozidla pozorovaná ve všech z určených pozic  $\alpha$ .

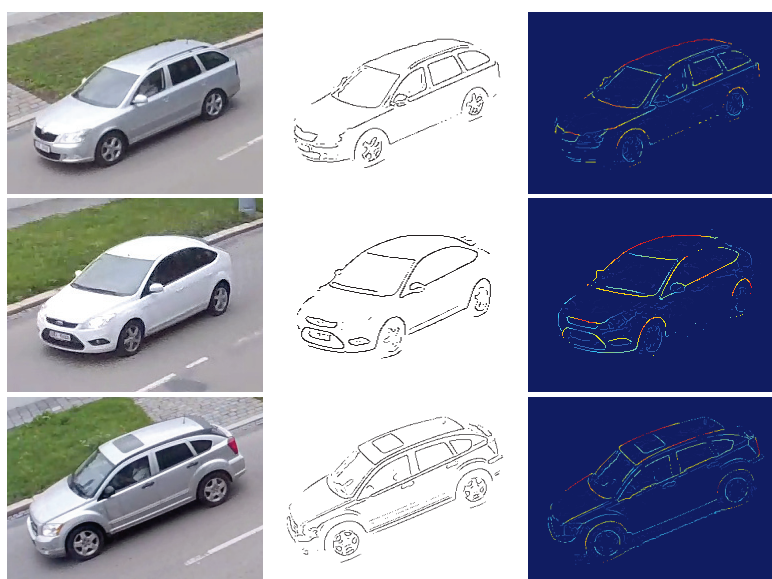


Obrázek 5.2: Akvizice vzorků vozidel. Dohledová kamera sleduje pohyblivá vozidla (zelená čára vyznačuje jednu trajektorii vozidla). Několik pozic v obraze (typicky 3) je manuálně vyznačeno (červené křížky) a pozice vozidla, které s nimi korespondují, jsou vybrány na základě pozice středu podstavy jejich obalového kvádrů (kapitola 4.2).

## 5.2 Extrakce příznaků pro shlukování vzorků vozidel

Naším cílem je zaznamenat vzorky každého konkrétního modelu vozidla za různých světelných podmínek, výskytu stínů, obsahující různé odlesky, apod. Jednou výraznou charakteristikou, jež je stabilní napříč působení těchto faktorů, jsou dominantní hrany [71], zvláště když je pro srovnání k dispozici více než jeden pohled na dané vozidlo. V takovém případě různé siluety a struktury hran implicitně reflektují trojrozměrný tvar vozidla [71]. Detekci hran provádíme detektorem hran Canny [84] a pomocí Structured Edge Detector, SED [80], viz obrázek 5.3. Pro další zpracování jsou hrany maskovány konvexní obálkou popředí vozidla tak, aby zpracovávány byly pouze hrany relevantní pro posouzení vozidla.

Rysy jednoho vzorku vozidla pro jednu pozici  $\alpha$  jsou reprezentovány jako  $\mathcal{S}^\alpha = (\mathcal{P}^\alpha, r^\alpha)$ ,



Obrázek 5.3: Extrakce charakteristických hran/kontur. **vlevo:** vstupní obraz, **uprostřed:** hrany získané detektorem Canny [84], **vpravo:** hrany získané detektorem Structured Edge Detector [80].

kde  $\mathcal{P}^\alpha = \{(x_1, y_1), \dots, (x_n, y_n)\}$  je množina souřadnic hranových bodů a  $r^\alpha : \mathcal{P}^\alpha \mapsto \langle 0, 1 \rangle$  je funkce zobrazující každou hranu na odpovídající odezvu detektoru SED. Úplná reprezentace jednoho vzorku vozidla je  $\mathcal{S} = \{\mathcal{S}^\alpha \mid \forall \alpha \in \mathbb{P}\}$ .

### 5.3 Cenová funkce rozdílnosti vozidel

Pro každý pár obrázků, které mají být porovnány, jsou nejprve hrany zarovnány pomocí Iterative Closest Points, ICP [81], viz obrázek 5.4. Dovolujeme, aby hranové body testovaného vozidla byly transformovány translací, rotací a uniformní změnou měřítka. Úvodní zarovnání pro ICP je získáno zarovnáním dvourozměrných těžišť obrazů vozidel. Během výpočtu ICP jsou využívány odezvy detektoru SED  $r^\alpha$  – registrační chyba je počítána s váhou, která je dána odezvou SED. Poté, co referenční vozidlo  $i$  a srovnávané vozidlo  $j$  na pozici  $\alpha$  jsou zarovnány transformací  $\phi(p)$  nalezenou pomocí ICP, cena rozdílnosti  $d_{ij}^\alpha$  vzorků je vypočtena takto:

$$d_{ij}^\alpha = \frac{|\{p \in \mathcal{P}_j^\alpha \mid \delta_{\min}(\phi(p), \mathcal{P}_i^\alpha) > \lambda\}|}{|\mathcal{P}_j^\alpha|} \quad (5.1)$$

$$\delta_{\min}(p, \mathcal{P}) = \min_{p' \in \mathcal{P}} \|p - p'\|. \quad (5.2)$$

Cena rozdílnosti (5.1) je vypočtena jako podíl hranových pixelů vozidla bez blízké shody s hranovými pixely referenčního vozidla. Pro větší přesnost a robustnost jsou do výsledné ceny rozdílnosti zahrnuty příspěvky ze všech určených pozic  $\alpha$ . Výsledná cena  $d_{ij}$  pro dva úplné vzorky vozidel je spočtena jako

$$d_{ij} = \max_{\alpha \in \mathbb{P}} [\max(d_{ij}^\alpha, d_{ji}^\alpha)], \quad (5.3)$$

kde  $\mathbb{P}$  je množina určených pozic vozidel (obrázek 5.2). Stojí za povšimnutí, že cena rozdílnosti je symetrická (tj.  $d_{ij} = d_{ji}$ ) z podstaty výpočtu; její symetričnost je důležitá pro algoritmus shlukování.



Obrázek 5.4: Cena rozdílnosti mezi vzorky vozidel. **vlevo:** referenční vozidlo, **uprostřed:** srovnávané vozidlo, **vpravo:** vizualizace rozdílnosti hran (černě: referenční vozidlo, zeleně: pixely srovnávaného vozidla se shodou, červeně: pixely srovnávaného vozidla bez shody).



## 5.4 Automatické shlukování vzorků aut

Označme množinu vzorků vozidel, které mají být shlukovány, symbolem  $V$ . S použitím cen rozdílnosti je zkonstruován neorientovaný graf  $\mathcal{G}_\tau = (\mathcal{V}_\tau, \mathcal{E}_\tau)$  a v něm jsou shluky aut reprezentovány jako spojitě komponenty:

$$\mathcal{V}_\tau = \{v \in V \mid \exists u \in V : d_{uv} \leq \tau \wedge u \neq v\} \quad (5.4)$$

$$\mathcal{E}_\tau = \{(u, v) \in V^2 \mid d_{uv} \leq \tau \wedge u \neq v\}. \quad (5.5)$$

Graf a výsledné shluky jsou parametrizovány prahem  $\tau$ . Změnami tohoto prahu je možné vytvořit celou stromovou hierarchii shluků. Označme množinu spojitých komponent grafu  $\mathcal{G}_\tau$  symbolem  $\mathcal{C}_\tau$  a předpokládejme, že práh  $\tau$  je vzorkován s malým krokem  $\varepsilon$  (např.  $\varepsilon = 0.01$ ). Stromová hierarchická struktura je reprezentována grafem  $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$ :

$$\mathcal{V}_T = \{(c, \tau) \mid \tau \in \{0, \varepsilon, 2\varepsilon, \dots, 1\} \wedge c \in \mathcal{C}_\tau\} \quad (5.6)$$

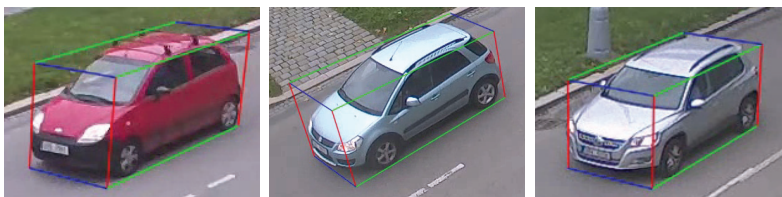
$$\mathcal{E}_T = \{((c_1, \tau_1), (c_2, \tau_2)) \in \mathcal{V}_T^2 \mid c_1 \supseteq c_2 \wedge \tau_1 > \tau_2\} \quad (5.7)$$

Příklad takové stromové hierarchie je na obrázku 5.6.

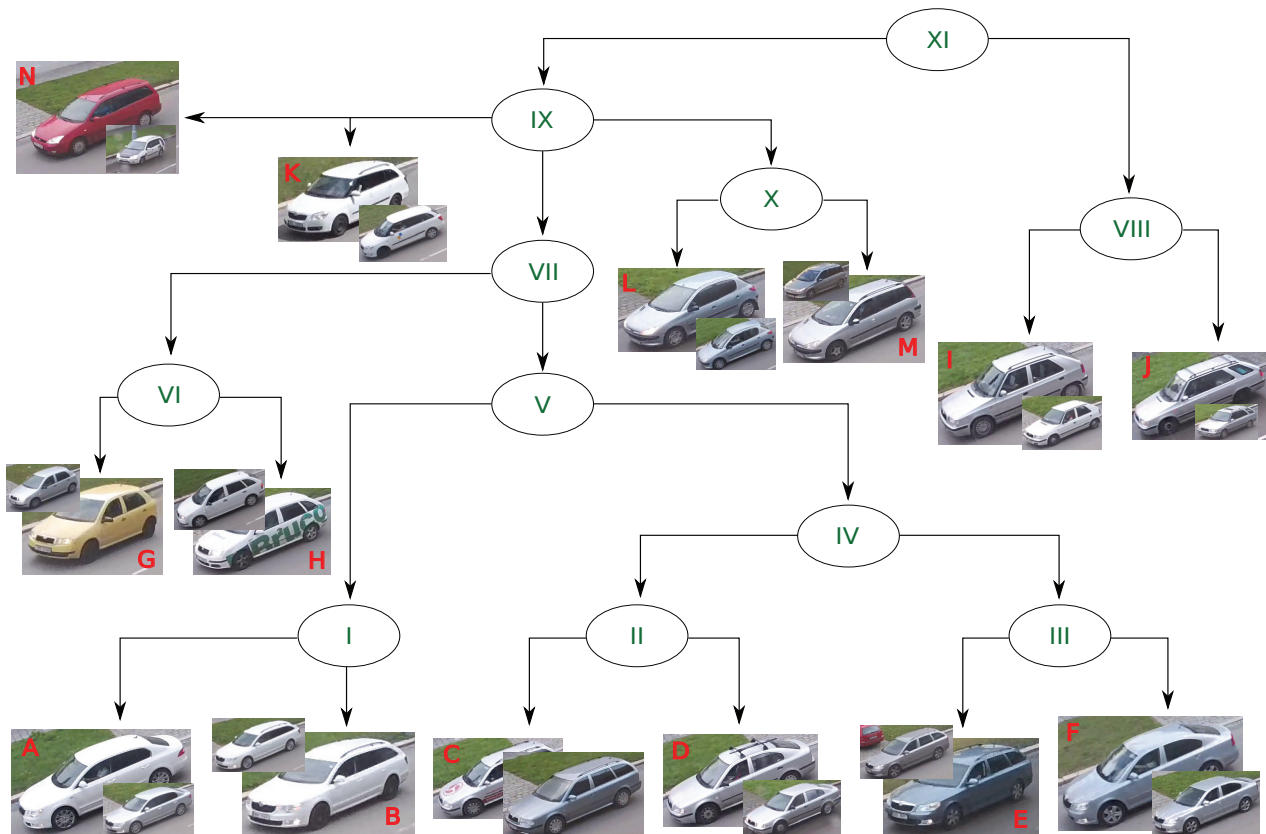
## 5.5 Vzájemná registrace shluků vozidel

Jak bylo uvedeno dříve (kapitola 4.2), pro každé vozidlo, které statická dohledová kamera pozoruje, je možné plně automaticky zkonstruovat trojrozměrný obalový kvádr. Příklady takto zkonstruovaných kvádrů pro auta z řešené datové sady ukazuje obrázek 5.5. Rozměry získané pro obalové kvádry společně s dalšími vlastnostmi shluků (např. relativní velikost shluku, jeho řád, atd.) mohou být použity pro vyhledávání obdobných shluků (k jejich vzájemné registraci). Je možné a smysluplné registrovat shluky pořízené v různé časové okamžiky jednou kamerou a nebo shluky získané rozdílnými kamerami. Každý shluk v hierarchii  $(c, \tau)$  je charakterizován příznakovým vektorem  $f : \mathcal{V}_T \mapsto \mathbb{R}^n$  složeným ze zmiňovaných příznaků. Stromy konstruované pro různé časové periody v rámci jedné kamery a nebo konstruované pro dvě různé kamery jsou registrovány existujícími algoritmy registrace grafů [85, 86].

Registrace shluků umožňuje množství nových případů užití. Ve scénáři shromažďování datové sady vozidel jsou anotace pořízené lidským operátorem nad jedním stromem přenášeny na jiný strom, pořízený z dat jiné kamery (nebo více jiných kamer). Je tak možné dosáhnout spolehlivých anotací výrobce a modelu vozidla pro data pozorovaná z různých úhlů pohledu. Další slibnou aplikací registrace shluků je přesná kalibrace měřítka kamery. V kapitole 4.3 je použit pro určení měřítka scény a kalibraci kamery medián rozměrů vozidel v každém směru. Mediány počítané individuálně pro větší počet konkrétních typů vozidel umožní výrazně přesnější a spolehlivější kalibraci a zároveň poskytnou vodítko pro odhad spolehlivosti provedené kalibrace.



Obrázek 5.5: Příklady trojrozměrných obalových kvádrů zkonstruovaných podle kapitoly 4.2. Délky, šířky a výšky jsou označeny zelenou, modrou a červenou barvou.



Obrázek 5.6: Dendrogram zobrazující část stromu vzniklého automatickým shlukováním (10k vzorků vozidel pozorovaných jednou kamerou). Typy vozidel v listech: a) Superb sedan, b) Superb combi, c) Octavia Mk1 combi, d) Octavia Mk1 sedan, e) Octavia Mk2 combi, f) Octavia Mk2 sedan, g) Fabia Mk1 hatchback, h) Fabia Mk1 combi, i) Favorit, j) Favorit combi, k) Fabia Mk2 combi, l) P206 hatchback, m) P206 combi, n) Focus combi. Jak je z dendrogramu patrné, vozidla jsou automaticky shlukována sémanticky smysluplným způsobem a shluky tvoří vhodný strom. Vnitřní uzly reprezentují například: I) Superb sedan + Superb combi, II) Octavia Mk1 combi + Octavia Mk1 sedan, III) Octavia Mk2 combi + Octavia Mk2 sedan, IV) všechny varianty vozu Octavia, V) všechny vozy střední velikosti a větší značky Škoda.

## 5.6 Předběžné výsledky

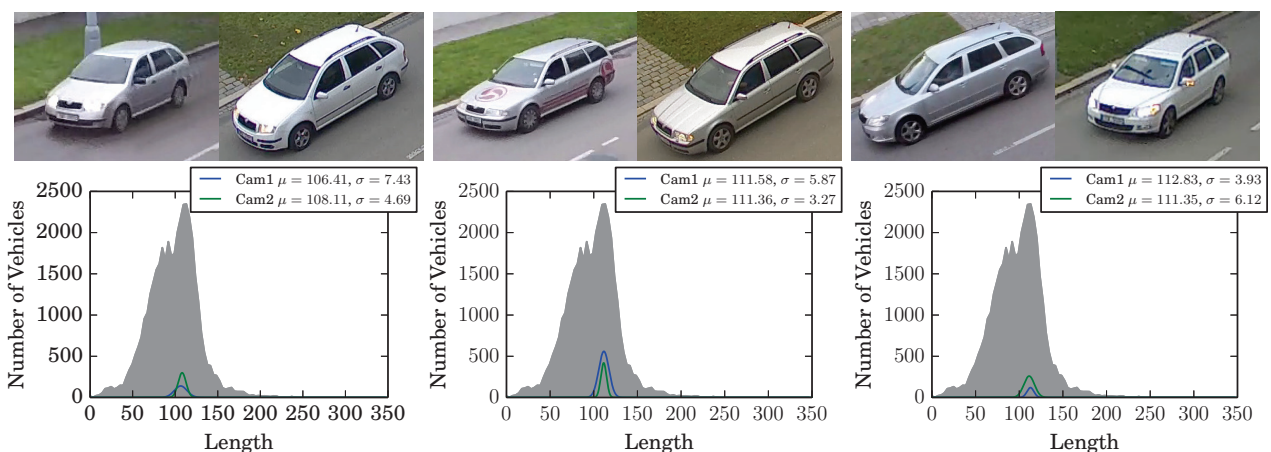
Celá kapitola 5 popisuje dosud nepublikovanou práci (v procesu přípravy k publikaci). Zde uvedeme výběr předběžných experimentálních výsledků, které naznačují potenciál metody. Delší dobu jsme sbírali vzorky vozidel kamerou pozorující málo frekventovanou ulici (Božetěchova ulice v Brně, obrázek 5.2). Kamera je poměrně standardní internetová kamera (IP camera) s nastavenou výraznou kompresí h264 a silným radiálním zkreslením – *AXIS M1034-W Network Camera*. Na 10 000 vzorcích (shromážděných během přibližně 50 dní) jsme provedli několik experimentů se shlukováním, abychom poznali vlastnosti navrhovaného algoritmu. Obrázek 5.7 ukazuje koláž vzorků jednoho listového uzlu stromu získaného při  $\tau = 0.38$ , což je dosti malý práh a shluk je „úzký“ v tom smyslu, že auta jsou si výrazně podobná a ve shluku není přítomno *žádné* auto, které by do něho nepatřilo.

Kromě 10 000 obrázků zmiňovaných výše jsme získali další sadu 10 000 obrázků z druhé kamery pozorující stejnou scénu z poněkud odlišného úhlu pohledu a v jiném čase. Použili jsme shlukovací procedury popsané v této kapitole na obě sady obrázků a registrovali jsme uzly

stromů vzniklých pro obě kamery. Obrázek 5.8 ilustruje dosažené výsledky: zobrazuje histogramy délek obalových kvádrů vozidel (délka je nejdiskriminativnější ze tří rozměrů obalových těles díky přirozenému vzhledu automobilů a díky tomu, že je kamery zabírají dostatečně z boku) a jejich podíl na celé sadě vzorků vozidel pozorovaných na silnici.



Obrázek 5.7: Příklad jednoho shluku získaného navrhovaným algoritmem. Shluk byl získán při  $\tau = 0.38$ , což znamená, že vzorky vozidel jsou si výrazně vizuálně podobné – například mají pouze světlé barvy (šedou, stříbrnou, bílou). Pro vyšší hodnoty  $\tau$  obsahují uzly větší výběr barev (jako v obrázku 5.1). Opakované pozorování stejného vozidla je způsobeno tím, že stejné vozidlo projelo silnicí vícekrát.



Obrázek 5.8: Shluky registrované mezi dvěma odlišnými kamerami pozorujícími stejnou scénu z jiných pohledů a v jiném čase. **horní řádek:** Příklady párů vozidel pozorovaných ze dvou kamer. **dolní řádek:** Histogramy délek (po registraci měřítka); barevně jsou histogramy délek vozidel v jednotlivých shlucích (tři největší shluky,  $\tau = 0.37$ ), šedě je histogram délek kompletní množiny vozidel pozorovaných v první kamerě.

## 6 Závěr

Tento text shrnuje několik prací autora a jeho spolupracovníků. Vybrány do něj byly nedávné práce směřující k plně automatickému zpracování obrazu z dohledových dopravních kamer. Cílem popisovaného výzkumu je odstranit požadavky na manuální kalibraci a nastavování jednotlivých kamer lidským operátorem. To je důležité pro usnadnění a automatizaci nasazení dohledu a prediktivní analýzy dopravních cest.

Všechny popisované metody staví na geometrické parametrizaci přímk [14], která umožňuje jejich detekci v digitálním obraze. Nad touto parametrizací je postavena její kaskádovaná varianta [32], která jde dál v tom smyslu, že umožňuje přímou detekci úběžníků v obraze. Úběžníky jsou přitom kritickou charakteristikou některých trojrozměrných scén. Jedná se jednak o tzv. Manhattanské světy, ve kterých se vyskytují hrany (okna, hrany budov, římsy, apod.) dominantně ve třech navzájem kolmých směrech. Ve svých pracích ale ukazujeme, že třemi ortogonálními úběžníky je možné charakterizovat i proud přímočaře se pohybujících vozidel. Detekce těchto tří úběžníků umožňuje kalibraci kamery [59] a také konstrukci trojrozměrných obalových kvádrů [63], jejichž analýzou po vhodné dobu je možné určit měřítko kamery.

Popisované metody mají širokou škálu aplikací. V automobilové dopravě se kamerový dohled využívá velice intenzivně. Ve velké většině případů kamery pozorující dopravu ovšem nejsou využívány ke strojovému zpracování dat, ale pouze lidské obsluze umožňují nahlédnout na danou dopravní situaci. Většímu nasazení strojového zpracování dosud bránilo především to, že nastavení jedné kamery vyžadovalo značnou odbornou činnost lidského operátora, případně odměřování kalibračních vzdáleností přímo ve fyzické scéně. Vytvořením a publikováním popisovaných postupů autoři přispěli k větší automatizaci monitorování dopravy, která může být využita ke zvýšení bezpečnosti a snížení ekologické zátěže dopravy, jež je přese všechno úsilí stále značná.

# Literatura

- [1] Paul V. C. Hough. Method and means for recognizing complex patterns, Dec 1962. U.S. Patent 3,069,654.
- [2] D. H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. pages 714–725, 1987.
- [3] J. Princen, J. Illingworth, and J. Kittler. Hypothesis testing: A framework for analyzing and optimizing Hough transform performance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(4):329–341, 1994.
- [4] Prabir Bhattacharya, Azriel Rosenfeld, and Isaac Weiss. Point-to-line mappings as Hough transforms. *Pattern Recognition Letters*, 23(14):1705–1710, 2002.
- [5] Tinne Tuytelaars, Marc Proesmans, Luc Van Gool, and Esat Mi. The cascaded hough transform. In *In Proceedings of ICIP*, pages 736–739, 1997.
- [6] Tinne Tuytelaars, Marc Proesmans, and Luc Van Gool. The cascaded hough transform as support for grouping and finding vanishing points and lines. In *Proceedings of the International Workshop on Algebraic Frames for the Perception-Action Cycle*, AFPAC '97, pages 278–289, London, UK, 1997. Springer-Verlag.
- [7] Richard O. Duda and Peter E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15, 1972.
- [8] Adel A. Sewisy. Graphical techniques for detecting lines with the hough transform. *International Journal of Computer Mathematics*, 79(1):49–64, 2002.
- [9] R.S. Wallace. A modified Hough transform for lines. In *Proceedings of CVPR 1985*, pages 665–667, 1985.
- [10] Frank Natterer. *The mathematics of computerized tomography*. Wiley, John & Sons, Incorporated, 1986. ISBN 9780471909590.
- [11] Ulrich Eckhardt and Gerd Maderlechner. Application of the projected Hough transform in picture processing. In *Proceedings of the 4th International Conference on Pattern Recognition*, pages 370–379, London, UK, 1988. Springer-Verlag.
- [12] Arthur V. Forman. A modified Hough transform for detecting lines in digital imagery. In *Applications of artificial intelligence III*, pages 151–160, 1986.
- [13] Alfred Inselberg. *Parallel Coordinates; Visual Multidimensional Geometry and Its Applications*. Springer, 2009. ISBN: 978-0-387-21507-5.
- [14] Markéta Dubská, Adam Herout, and Jiří Havel. PClines – line detection using parallel coordinates. In *Proceedings of CVPR 2011*, 2011.
- [15] S. El Mejdani, R. Egli, and F. Dubeau. Old and new straight-line detectors: Description and comparison. *Pattern Recogn.*, 41:1845–1866, June 2008.

- [16] Patrick Denis, James H. Elder, and Francisco J. Estrada. Efficient edge-based methods for estimating manhattan frames in urban imagery. In *Proc. ECCV*, 2008.
- [17] Matthew E. Antone and Seth Teller. Automatic recovery of relative camera rotations for urban scenes. In *Proc. CVPR*, 2000.
- [18] Ciarán Hughes, Patrick Denny, Martin E. Glavin, and Edward Jones. Equidistant fish-eye calibration and rectification by vanishing point extraction. *PAMI*, 2010.
- [19] R. Cipolla, T. Drummond, and D. Robertson. Camera calibration from vanishing points in images of architectural scenes. In *Proc. BMVC*, 1999.
- [20] Jana Košecká and Wei Zhang. Video compass. In *ECCV*. 2002.
- [21] Jefferey A. Shufelt. Performance evaluation and analysis of vanishing point detection techniques. *IEEE Tran. Pattern Analysis and Machine Intelligence*, 21:282–288, 1999.
- [22] Stephen T. Barnard. Interpreting perspective images. *Artif. Intell.*, 21(4):435–462, November 1983.
- [23] Tinne Tuytelaars, Marc Proesmans, Luc Van Gool, and Esat Mi. The cascaded Hough transform. In *Proceedings of ICIP 1998*, pages 736–739, 1998.
- [24] Bo Li, Kun Peng, Xianghua Ying, and Hongbin Zha. Vanishing point detection using cascaded 1D hough transform from single images. *Pattern Recognition Letters*, 2012.
- [25] J.-C. Bazin, C. Demonceaux, P. Vasseur, and I. Kweon. Rotation estimation and vanishing point extraction by omnidirectional vision in urban environment. *IJRR*, 31, 2012.
- [26] Jean-Charles Bazin and Marc Pollefeys. 3-line RANSAC for orthogonal vanishing point detection. In *Proc. IROS*, 2012.
- [27] Wolfgang Förstner. Optimal vanishing point detection and rotation estimation of single images from a legoland scene. In *ISPRS*, 2010.
- [28] F.M. Mirzaei and S.I. Roumeliotis. Optimal estimation of vanishing points in a manhattan world. In *ICCV*, 2011.
- [29] Marcos Nieto and Luis Salgado. Real-time robust estimation of vanishing points through nonlinear optimization. In *RTIVP*, 2010.
- [30] W. Elloumi, S. Treuillet, and R. Leconge. Tracking orthogonal vanishing points in video sequences for a reliable camera orientation in manhattan world. In *CISP*, 2012.
- [31] Marcos Nieto and Luis Salgado. Simultaneous estimation of vanishing points and their converging lines using the em algorithm. *Pattern Recognition Letters*, 2011.
- [32] Markéta Dubská and Adam Herout. Real projective plane mapping for detection of orthogonal vanishing points. In *British Machine Vision Conference, BMVC*, 2013.
- [33] Jiří Havel, Adam Herout, and Markéta Dubská. Vanishing points in point-to-line mappings and other line parameterizations. *Pattern Recognition Letters*, 34:703–708, 2013.
- [34] J de Ortuzar and Luis G Willumsen. *Modelling transport*. 2011.

- [35] Stefan Lämmer and Dirk Helbing. Self-control of traffic lights and vehicle flows in urban road networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(04), 2008.
- [36] Shunsuke Kamijo, Yasuyuki Matsushita, Katsushi Ikeuchi, and Masao Sakauchi. Traffic monitoring and accident detection at intersections. *Intelligent Transportation Systems, IEEE Transactions on*, 1(2):108–118, 2000.
- [37] David Vallejo, Javier Albusac, Luis Jimenez, Carlos Gonzalez, and Juan Moreno. A cognitive surveillance system for detecting incorrect traffic behaviors. *Expert Systems with Applications*, 36(7):10503–10511, 2009.
- [38] Lazaros Grammatikopoulos, George Karras, and Elli Petsa. Automatic estimation of vehicle speed from uncalibrated video sequences. In *Proceedings of International Symposium on Modern Technologies, Education and Professional Practice in Geodesy and Related Fields*, pages 332–338, 2005.
- [39] Kunfeng Wang, Hua Huang, Yuantao Li, and Fei-Yue Wang. Research on lane-marking line based camera calibration. In *International Conference on Vehicular Electronics and Safety, ICVES*, 2007.
- [40] F.W. Cathey and D.J. Dailey. A novel technique to dynamically measure vehicle speed using uncalibrated roadway cameras. In *Intelligent Vehicles Symposium*, pages 777–782, 2005.
- [41] Tun-Wen Pai, Wen-Jung Juang, and Lee-Jyi Wang. An adaptive windowing prediction algorithm for vehicle speed estimation. In *IEEE Intelligent Transportation Systems*, 2001.
- [42] D.J. Dailey, F.W. Cathey, and S. Pumrin. An algorithm to estimate mean traffic speed using uncalibrated cameras. *IEEE Transactions on Intelligent Transportation Systems*, 1(2):98–107, 2000.
- [43] Tuan Hue Thi, Sijun Lu, and Jian Zhang. Self-calibration of traffic surveillance camera using motion tracking. In *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems*, 2008.
- [44] T.N. Schoepflin and D.J. Dailey. Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation. *IEEE Transactions on Intelligent Transportation Systems*, 4(2):90–98, 2003.
- [45] Bruno Caprile and Vincent Torre. Using vanishing points for camera calibration. *International Journal of Computer Vision*, 4(2):127–139, 1990.
- [46] J. Deutscher, M. Isard, and J. MacCormick. Automatic camera calibration from a single manhattan image. In *European Conference on Computer Vision, ECCV*, pages 175–188, 2002.
- [47] Kai-Tai Song and Jen-Chao Tai. Dynamic calibration of Pan-Tilt-Zoom cameras for traffic monitoring. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 36(5):1091–1103, 2006.
- [48] Rong Dong, Bo Li, and Qi-mei Chen. An automatic calibration method for PTZ camera in expressway monitoring system. In *World Congress on Computer Science and Information Engineering*, pages 636–640, 2009.

- [49] Yuan Zheng and Silong Peng. A practical roadside camera calibration method based on least squares optimization. *Intelligent Transportation Systems, IEEE Transactions on*, 15(2):831–843, April 2014.
- [50] George S. K. Fung, Nelson H. C. Yung, and Grantham K. H. Pang. Camera calibration from road lane markings. *Optical Engineering*, 42(10):2967–2977, 2003.
- [51] Bart De Schutter and Bart De Moor. Optimal traffic light control for a single intersection. *European Journal of Control*, 4(3):260–276, 1998.
- [52] Neeraj K Kanhere, Stanley T Birchfield, and Wayne A Sarasua. Automatic camera calibration using pattern detection for vision-based speed sensing. *Journal of the Transportation Research Board*, 2086(1):30–39, 2008.
- [53] Zhaoxiang Zhang, Tieniu Tan, Kaiqi Huang, and Yunhong Wang. Practical camera calibration from moving objects for traffic scene surveillance. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(3):518–533, 2013.
- [54] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik. A real-time computer vision system for measuring traffic parameters. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 1997.
- [55] N. K. Kanhere and S. T. Birchfield. Real-time incremental segmentation and tracking of vehicles at low camera angles using stable features. *IEEE Transactions on Intelligent Transportation Systems*, 9(1):148–160, 2008.
- [56] Xiao Chen He and N. H C Yung. A novel algorithm for estimating vehicle speed from two consecutive images. In *IEEE Workshop on Applications of Computer Vision, WACV*, 2007.
- [57] F.W. Cathey and D.J. Dailey. Mathematical theory of image straightening with applications to camera calibration. In *Intelligent Transportation Systems Conference*, 2006.
- [58] Zhigang Zhu, Bo Yang, Guangyou Xu, and Dingji Shi. A real-time vision system for automatic traffic monitoring based on 2D spatio-temporal images. In *Proceedings of WACV*, 1996.
- [59] Markéta Dubská, Adam Herout, Jakub Sochor, and Roman Juránek. Fully automatic roadside camera calibration for traffic surveillance. *Accepted for publication in IEEE Transactions on Intelligent Transportation Systems*.
- [60] C. Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition*, volume 2, pages 246–252, 1999.
- [61] Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 2, pages 28–31 Vol.2, 2004.
- [62] T. Horprasert, D. Harwood, and L. S. Davis. A statistical approach for real-time robust background subtraction and shadow detection. In *Proc. IEEE ICCV*, volume 99, pages 1–19, 1999.



- [63] Markéta Dubská, Jakub Sochor, and Adam Herout. Automatic camera calibration for traffic understanding. In *Proceedings of the British Machine Vision Conference (BMVC)*, 2014.
- [64] J. M. Wang, Y. C. Chung, C. L. Chang, and S.W. Chen. Shadow detection and removal for traffic images. In *Networking, Sensing and Control, 2004 IEEE International Conference on*, volume 1, pages 649–654 Vol.1, March 2004.
- [65] Mei Xiao, Chong-Zhao Han, and Lei Zhang. Moving shadow detection and removal for traffic sequences. *International Journal of Automation and Computing*, 4(1):38–46, 2007.
- [66] V.S. Petrovic and T. F. Cootes. Analysis of features for rigid structure vehicle type recognition. In *Proceedings of the British Machine Vision Conference (BMVC)*, pages 587–596, 2004.
- [67] Louka Dlagnekov and Serge Belongie. Recognizing cars. Technical report, UCSD CSE Tech Report CS2005-0833, 2005.
- [68] G. Pearce and N. Pears. Automatic make and model recognition from frontal images of cars. In *IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, pages 373–378, 2011.
- [69] P. Negri, X. Clady, M. Milgram, and R. Poulencard. An oriented-contour point based voting algorithm for vehicle type classification. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 1, pages 574–577, 2006.
- [70] Bailing Zhang. Classification and identification of vehicle type and make by cortex-like image descriptor hmax. *International Journal of Computational Vision and Robotics*, 4:195–211, 2014.
- [71] E. Hsiao, S.N. Sinha, K. Ramnath, S. Baker, L. Zitnick, and R. Szeliski. Car make and model recognition using 3D curve alignment. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–1, March 2014.
- [72] Michael Stark, Jonathan Krause, Bojan Pepik, David Meger, James Little, Bernt Schiele, and Daphne Koller. Fine-grained categorization for 3D scene understanding. In *Proceedings of the British Machine Vision Conference (BMVC)*, 2012.
- [73] Jonathan Krause, Jia Deng, Michael Stark, and Li Fei-Fei. Collecting a large-scale dataset of fine-grained cars. In *Second Workshop on Fine-Grained Visual Categorization (FGVC<sup>2</sup>)*, 2013.
- [74] Yen-Liang Lin, Vlad I. Morariu, Winston Hsu, and Larry S. Davis. Jointly optimizing 3D model fitting and fine-grained classification. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2014.
- [75] Jonathan Krause, Michael Stark, Jia Deng, and Li Fei-Fei. 3D object representations for fine-grained categorization. In *IEEE Workshop on 3D Representation and Recognition, at ICCV 2013 (3dRR-13)*, 2013.
- [76] Jinjun Wang, Jianchao Yang, Kai Yu, Fengjun Lv, T. Huang, and Yihong Gong. Locality-constrained linear coding for image classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3360–3367, June 2010.

- [77] Jia Deng, J. Krause, and Li Fei-Fei. Fine-grained crowdsourcing for fine-grained recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 580–587, June 2013.
- [78] N.S. Thakoor and B. Bhanu. Structural signatures for passenger vehicle classification in video. *IEEE Transactions on Intelligent Transportation Systems*, 14(4):1796–1805, Dec 2013.
- [79] J. Prokaj and G. Medioni. 3-d model based vehicle recognition. In *Applications of Computer Vision (WACV), 2009 Workshop on*, pages 1–7, Dec 2009.
- [80] P. Dollár and C.L. Zitnick. Structured forests for fast edge detection. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 1841–1848, Dec 2013.
- [81] P.J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2):239–256, Feb 1992.
- [82] C. Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition*, volume 2, pages 246–252, 1999.
- [83] Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 2, pages 28–31 Vol.2, 2004.
- [84] John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679–698, Nov 1986.
- [85] Feng Zhou and F. De La Torre. Deformable graph matching. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2922–2929, June 2013.
- [86] Yin-Tang Dai and Shi-Yong Zhang. A fast labeled graph matching algorithm based on edge matching and guided by search route. In *Wavelet Analysis and Pattern Recognition, 2009. ICWAPR 2009. International Conference on*, pages 1–7, July 2009.

## Abstrakt

Práce shrnuje autorův výzkum v oblasti automatického zpracování videozáběrů z dohledových kamer v dopravě. Většímu rozšíření automatického zpracování dopravních dat stále brání nutnost investovat nezanedbatelné expertní úsilí do nastavení každé jednotlivé monitorovací kamery. Přínosem popisovaných technik je zejména to, že vedou k odstranění tohoto omezení. V textu je popsána parametrizace PClines, která umožňuje efektivní detekci přímk v obraze. Na této parametrizaci staví kaskádovaná varianta „diamond space“, používaná k přímé detekci úběžníků, zejména trojic vzájemně ortogonálních úběžníků. Trojice ortogonálních oběžníků může popisovat dopravní proud vozidel a pozorováním proudu vozidel je možné skrze určení úběžníků provést kalibraci kamery, včetně určení měřítka. Tato kalibrace umožňuje měřit vzdálenosti a rychlost v jednotkách reálného světa. V poslední kapitole je popsán aktuální a probíhající výzkum autora a jeho kolegů v oblasti rozpoznání přesného typu vozidla, opět s výrazným snížením požadavků na ruční vstup.

## Abstract

This work summarizes author's research in the field of automatic processing of video streams from traffic surveillance cameras. Larger use of automatic processing of such data is still prevented by the need of investing considerable amount of human labor into setting up each individual monitoring camera. The contribution of the presented techniques is especially the fact that they lead towards removal of this hindrance. The text describes parameterization PClines which allows for effective detection of lines in raster images. Based on this parameterization, a cascaded version "diamond space" is built, which is used for direct detection of vanishing points, namely triplets of mutually orthogonal vanishing points. A triplet of orthogonal vanishing points can be used for describing the flow of vehicles in traffic, and by observing such a flow, it is possible to use the detected vanishing points for camera calibration, including the scale. This calibration allows for measuring distances and speed in the real-world units. The last chapter describes ongoing research of the author and his colleagues in the field of recognition of exact make & model of vehicles, again with a considerable reduction of the requirements on manual input.