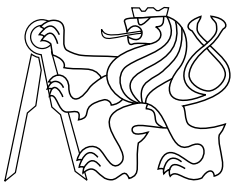
CENTER FOR
MACHINE PERCEPTION



CZECH TECHNICAL
UNIVERSITY IN PRAGUE

BACHELOR THESIS

ISSN 1213-2365

Planární segmentace sekvence hloubkových map

Miroslav Voda

vodamiro@gmail.com

CTU–CMP–2013–13

May 24, 2013

Available at
http://cmp.felk.cvut.cz/pub/cmp/articles/matousek/Voda-TR-2013-13.pdf

Thesis Advisor: Ing. Martin Matoušek, Ph.D.

Research Reports of CMP, Czech Technical University in Prague, No. 13, 2013

Published by

Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Technická 2, 166 27 Prague 6, Czech Republic
fax +420 2 2435 7385, phone +420 2 2435 7637, www: http://cmp.felk.cvut.cz

České vysoké učení technické v Praze
Fakulta elektrotechnická

Katedra kybernetiky

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: Miroslav V o d a

Studijní program: Otevřená informatika (bakalářský)

Obor: Informatika a počítačové vědy

Název tématu: Planární segmentace sekvence hloubkových map

Pokyny pro vypracování:

1. Použijte hloubkové mapy poskytované existující implementací stereopárování pracující s dopředu orientovaným párem kamer umístěným na jedoucím vozidle.
2. Navrhněte a implementujte metodu segmentace hloubkové mapy na rovinu terénu a soubor planárních objektů.
3. Využijte sekvenční charakter dat pro stabilizaci výsledků v čase.
4. Metoda musí být robustní vůči chybám v disparitní mapě.
5. Metodu implementujte v jazyce Matlab a otestujte na sekvenci z reálného provozu.

Seznam odborné literatury:

[1] Pfeiffer, David; Franke, Uwe: Towards a Global Optimal Multi-Layer Stixel Representation of Dense 3D Data. In Proc. British Machine Vision Conference, 2011.
[2] Dobiáš, Martin; Šára, Radim: Real-Time Global Prediction for Temporally Stable Stereo. In Proc ICCV Workshop on Live Dense Reconstruction from Moving Cameras, November 2011.
[3] Havlena, Vladimír; Štecha, Jan: Moderní teorie řízení. Kapitola 7.3 Kalmanův filtr. Skriptum FEL, Vydavatelství ČVUT, 1996.

Vedoucí bakalářské práce: Ing. Martin Matoušek, Ph.D.

Platnost zadání: do konce zimního semestru 2013/2014

L.S.

prof. Ing. Vladimír Mařík, DrSc.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 10. 1. 2013

Poděkování

Rád bych poděkoval vedoucímu bakalářské práce Ing. Martinovi Matouškovi, Ph.D. za jeho ochotu, trpělivost a odbornou pomoc při vytváření této práce. Dále děkuji všem blízkým a přátelům za jejich podporu.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne 24. 5. 2013 .....

Abstract

The method for segmentation of a sequence of disparity maps is proposed and implemented in this thesis. The maps are segmented into ground plane and a set of planar objects. The disparity maps are obtained from a forward-looking stereo-camera placed on a moving vehicle. The proposed plane estimation method uses Kalman filter for temporal stabilization. The segmentation into a set of objects, using the previously estimated ground plane, is based on [6]. The revealed optimization problem is solved using dynamic programming. All implemented algorithms and processes are written in *Matlab* programming language and then evaluated on a sequence obtained from real traffic.

Abstrakt

V této práci je navržena a implementována metoda segmentace sekvence hloubkových map na rovinu terénu a soubor planárních objektů. Vstupními daty určenými k segmentaci jsou hloubkové mapy získané z dopředu orientovaného páru kamer, který je umístěn na jedoucím vozidle. Metoda hledání roviny terénu v hloubkových mapách je stabilizována v čase pomocí Kalmanova filtru. Použitá metoda segmentace hloubkových map s využitím odhadnuté roviny terénu je založena na práci [6]. Formulovaný optimalizační problém je vyřešen za využití dynamického programování. Implementační část bakalářské práce je realizována v programovacím jazyce aplikace *Matlab*. Metoda je otestována na sekvenci z reálného provozu.

Obsah

1 Úvod **1**

2 Disparitní mapa **2**
2.1 Stereokamera . . . 2
2.2 Disparita a dispartní mapa . . . 2
2.3 Triangulace . . . 3
2.4 Použité metody výpočtu disparitní mapy . . . 4
2.5 Výsledky . . . 4

3 Rovina terénu **6**
3.1 Rovina terénu . . . 6
3.2 Hledání roviny v množině bodů . . . 6
3.3 Stabilizace roviny pomocí Kalmanovy filtrace . . . 10
3.4 Odhad roviny terénu s omezující podmínkou na horizont . . . 10
3.5 Vztah mezi modelem roviny terénu a disparitní mapou . . . 11
3.6 Výsledky . . . 13

4 Segmentace disparitní mapy **16**
4.1 Charakteristika scény . . . 16
4.2 *Stixel*ová reprezentace scény . . . 16
4.3 Klasifikace *Stixel*u . . . 17
4.4 Pravděpodobnostní model . . . 18
4.5 Implementace metody segmentace . . . 21
4.6 Výsledky . . . 26
4.7 Diskuze . . . 27

5 Závěr **30**

A Hledání optimální cesty **32**

Kapitola 1

Úvod

S pokračujícím růstem silniční dopravy se také klade větší důraz na bezpečnost silničního provoz. Proto se moderní dopravní prostředky opatřují různými bezpečnostními systémy, které snižují riziko střetu s ostatními účastníky provozu, jako jsou například ostatní dopravní prostředky, chodci či zvěř. Mezi tyto sofistikované bezpečnostní systémy bezpochyby patří detekce překážek v trajektorii vozidla. Stavebním kamenem v detekci překážek je zpracování 3D dat charakterizujících scénu před vozidlem. Mezi nejdostupnější způsoby získání těchto dat patří hloubkové či disparitní mapy získané z páru digitálních kamer umístěného na vozidle, který zachycuje situaci odehrávající se před vozidlem. Disparitní mapy jsou ve většině případů zašuměné, poskytují velké množství informací a obsahují mnoho děr. Proto jsou pro samotné vyhodnocení situace nevhodné. Aby bylo možné dopravní situaci snadněji vyhodnotit, je potřeba informace obsažené v jednotlivých pixelech disparitní mapy převést na prvky charakterizující reálný svět. Pro detekci překážek v prostředí silničního provozu zejména na terén a ostatní objekty (překážky). Na základě takového rozdělení scény lze danou situaci vyhodnotit a dále zpracovat (např. upozornit řidiče na překážku nebo zasáhnout do řízení). Proto je v této práci navržen, implementován a vyhodnocen postup pro segmentaci disparitních map na rovinu terénu a soubor planárních objektů kolmých na směr pohledu.

V kapitole 2 je vysvětlen postup získání disparitní mapy. Kapitola 3 popisuje navrženou metodu odhadnutí roviny terénu z hloubkové mapy, její kriteriální ohodnocení a stabilizaci v čase. Odhad roviny je realizován za použití metody *RANSAC* na množině bodů extrahovaných z disparitní mapy, stabilizace je provedena pomocí Kalmanovy filtrace. V kapitole 4 je navržen a implementován proces segmentace vstupních dat s využitím odhadu roviny terénu a metody segmentace založené na *Stixel*ové reprezentaci a pravděpodobnostním modelu dle [1],[6]. Většina použitých algoritmů je implementována ve skriptovacím jazyce *Matlab*. V této kapitole je také zhodnocení dosažených výsledků za použití různých metod získání disparitní mapy a volbě optimalizačních parametrů algoritmu segmentace. V kapitole 5 je shrnutí celé práce.

Kapitola 2

Disparitní mapa

Tato kapitola obsahuje vysvětlení postupu získávání disparitních map a poskytuje výčet metod použitých pro jejich výpočet.

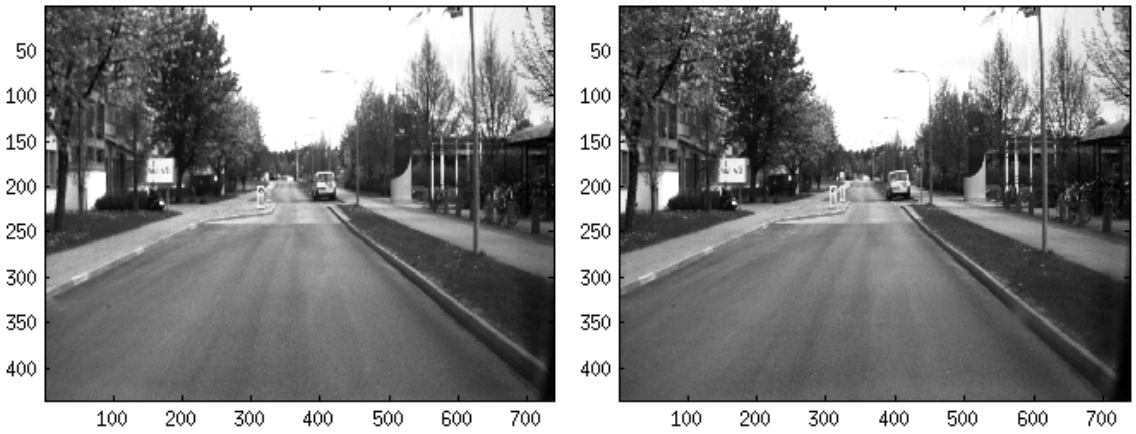
2.1 Stereokamera

Stereokamera je pár synchronizovaných kamer, který umožňuje pořízení epipolárně srovnané[1] [4] dvojice snímků (Obr. 2.1) scény ve stejném okamžiku za stejných podmínek a nastavení. Kamera použitá pro pořízení vstupních dat použitých v této práci je *PointGrey BumbleBee®2* [8] stereokamera, Obr. 2.2a.

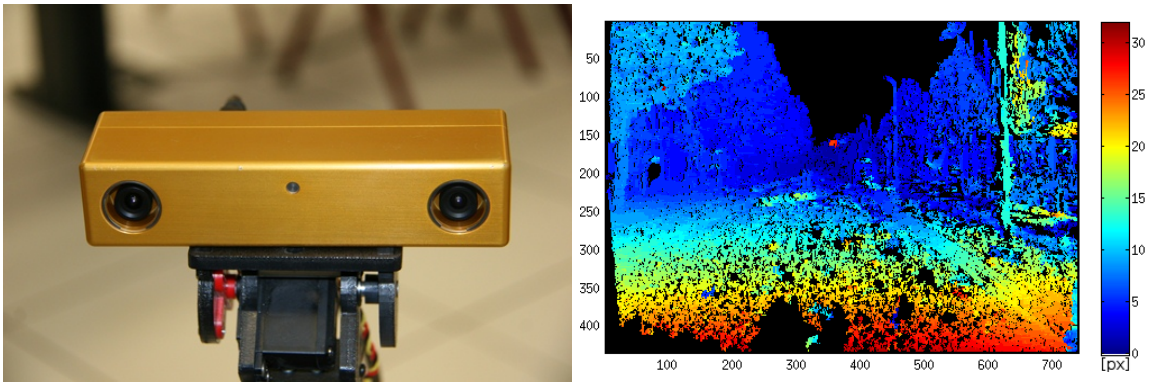
2.2 Disparita a dispartní mapa

Disparita je rozdíl x-ových souřadnic korespondujícího bodu na daném řádku y v levém a pravém snímku (Obr. 2.1) epipolárně srovnaného páru. Disparitní mapa (Obr. 2.2b) je

[1] Pokud snímky nejsou srovnané, je možné je následně srovnat geometrickou transformací



Obrázek 2.1: Ukázka páru fotografií



(a) (b)

Obrázek 2.2: BumbleBee stereokamera (a) a získaná disparitní mapa (b) z páru snímků (Obr. 2.1)

zobrazení $x \times y \to d$, kde x a y jsou souřadnice pixelu a d je disparita (v pixelech). Toto zobrazení lze reprezentovat maticí o velikosti snímku, kde hodnota prvků matice je rovna disparitě $d_{(x,y)}$. Pixel (x, y) v levém snímku tedy koresponduje s pixelem $(x + d_{(x,y)}, y)$ v pravém snímku.

2.3 Triangulace

Triangulace je metoda zjišťování souřadnic bodu pomocí trojúhelníku, kde dva referenční body (vrcholy trojúhelníku) jsou známy. Situaci zachycuje Obr. 2.3.

Mějme neznámý bod $P = [X, Y, Z]$ v prostoru a souřadnice jeho projekce na levý a pravý snímek. Projekce bodu P na levý snímek má v ortonormálním souřadném systému se středem v hlavním bodě[2] souřadnice $[x_L, y]$, projekce na snímek pravý $[x_R, y]$. Disparita d je tedy rovna rozdílu x-ových souřadnic x_L a x_R,

$$d = x_L - x_R \,. \tag{2.1}$$

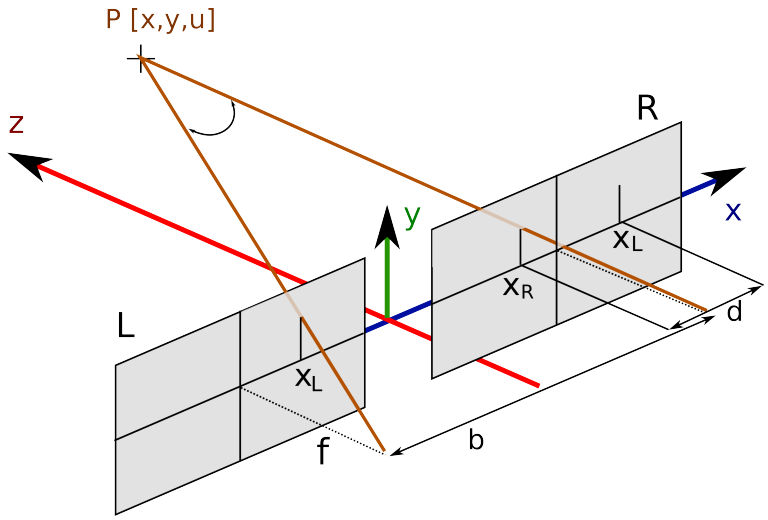
Souřadnice bodu P pak lze spočítat jako

$$X = \frac{bf}{x_L}, \quad Y = \frac{bf}{y}, \quad Z = \frac{bf}{d} \,, \tag{2.2}$$

kde b je délka základny[3], a f je ohnisková vzdálenost. Z tohoto vztahu je zřejmé, že čím se objekt ve scéně nachází dále, tím je disparita nižší (modrá barva - Obr. 2.2b), čím je objekt blíže (červená barva - Obr. 2.2b), tím je disparita vyšší. Černá barva značí díry v disparitní mapě. Objekt nacházející se ve scéně nekonečně daleko má disparitu rovnu nule.

[2] Hlavní bod je průsečík optické osy s obrazovou rovinou

[3] anglicky *baseline*



Obrázek 2.3: Projekce bodu na epipolárně srovnaný pár snímků

Popis	Přepínač	Parametr
Algoritmus	-alg	baseline |proposed
Korelátor	-corr	mncc |ncc |census |sad
Minimální disparita	-mindisp	%integer%
Maximální disparita	-maxdisp	%integer%
Vyplnění děr	-fillholes	
Práh	-thr	%float%
Počáteční práh	-initthr	%float%

Tabulka 2.1: Parametry rozšíření

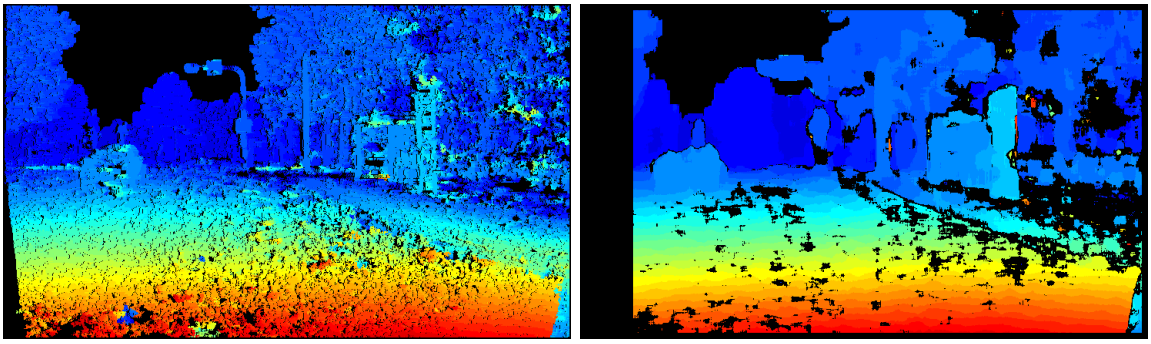
2.4 Použité metody výpočtu disparitní mapy

Growth Základní metodou pro výpočet disparitní mapy se stala aplikace *Growth* [2], pro kterou jsem za použití poskytnutých zdrojových kódů v jazyce *C++* naprogramoval *MEX* [10] rozšíření tvořící mezičlánek mezi aplikacemi *Matlab* a *Growth*. Přehled dostupných parametrů metody *Growth* nabízí Tab. 2.1.

Vestavěná *Matlab* metoda Jako druhý způsob získaní disparitní mapy je použita metoda *disparity* obsažená v *Computer Vision System Toolbox* aplikace *Matlab*. Tato metoda nabízí větší množství vstupních parametrů ovlivňující kvalitu a výpočet disparitní mapy. Použití této metody dokumentuje referenční příručka [9]. V této práci jsou použity obě metody a výsledky jsou srovnány.

2.5 Výsledky

Obr. 2.4a reprezentuje disparitní mapu získanou pomocí aplikace *Growh*, kdežto Obr. 2.4b je disparitní mapa získaná vestavěnou metodou *disparity* z *Computer Vision System Toolbox*



(a) (b)

Obrázek 2.4: Disparitní mapy získané pomocí různých metod. Disparitní mapa získaná pomocí metody *Growth* (a) a získaná pomocí vestavěné metody *disparity* (b).

aplikace *Matlab*. Černá barva odpovídá nenalezené korespondenci mezi body v páru snímků, takovým bodům říkáme díry. Aplikace *Growh* poskytuje celočíselné disparitní mapy, které jsou oproti vestavěné metodě *disparity* (dále jen vestavěná metoda) hustěji vyplněny (obsahují méně děr), ale také oproti vestavěné metodě obsahují více špatně odhadnutých hodnot disparity (šumu). Vestavěná metoda poskytuje mapy s desetinnými hodnotami disparity, obsahuje větší počet děr a menší množství šumu. Na rozdíl od metody *Growth* obsahuje disparitní mapa shluky disparity s podobnou disparitou. Aby mohla být disparitní mapa získaná vestavěnou funkcí použita v implementované metodě segmentace, pak musí být zaokrouhlena na celá čísla.

Kapitola 3

Rovina terénu

Tato kapitola popisuje definici roviny terénu, její metodu hledání v disparitní mapě pomocí algoritmu *RANSAC* a stabilizaci za použití Kalmanovy filtrace.

3.1 Rovina terénu

Obecná rovina v prostoru $\mathbb{R}^3$ je množina bodů $\mathbf{x}$ odpovídající předpisu

$$\mathbf{n}^T\mathbf{x} + d = 0\,, \tag{3.1}$$

kde $\mathbf{n}$ je normálový vektor roviny a d je posunutí v počátku. Euklidovská vzdálenost bodu x od roviny (3.1) je definována jako funkce

$$\mathrm{dist}_{\mathrm{d}}^{\mathbf{n}}(x) = \mathbf{n}^T\mathbf{x} + d\,, \tag{3.2}$$

kde pro $\mathrm{dist}_{\mathrm{d}}^{\mathbf{n}}(\mathbf{x}) = 0$ leží bod $\mathbf{x}$ v rovině. Pro $\mathrm{dist}_{\mathrm{d}}^{\mathbf{n}}(\mathbf{x}) < 0$ se bod nachází pod rovinou, opačně pro $\mathrm{dist}_{\mathrm{d}}^{\mathbf{n}}(\mathbf{x}) > 0$ se nachází nad rovinou. V reálném světě zpravidla terén neodpovídá rovině, ale jelikož vstupními daty jsou snímky ze silničního provozu, pak je hledanou rovinou terénu vozovka, jejíž body lze s vysokou přesností aproximovat pomocí roviny. Definujme tedy rovinu terénu v množině bodů prostoru $\mathbb{R}^3$ jako rovinu, která se v ideálním případě nachází co nejníže tak, že se pod rovinou nenacházejí žádné jiné body[1] a zároveň množina bodů ležících v rovině nebo v její blízkosti (dané prahem vzdálenosti) je maximální možná. Dále pro rovinu terénu jakožto vozovku platí úhlová omezení, které zabraňují odhadnutí roviny pod příliš vysokými úhly.

3.2 Hledání roviny v množině bodů

Při hledání roviny v prostoru $\mathbb{R}^3$ tvořeného body $\mathbb{P}$ existuje nekonečně mnoho možností umístění roviny, kde pouze nějaká podmnožina reprezentací splňuje požadavky roviny terénu zmíněných v sekci 3.1. Pro jednoznačné určení polohy roviny v prostoru $\mathbb{R}^3$ jsou potřeba tři

[1] tudíž všechny ostatní body se nacházejí nad nebo v rovině terénu

různé body. Proto pro nalezení požadované roviny generujeme hypotézy a ověřujeme je pomocí kriteriální funkce. Nejlepší hypotéza se hledá maximalizací kriteriální funkce. Zkoušet všechny možné kombinace trojic bodů by bylo časově velmi náročné, ale protože se v odhadnuté množině bodů $\mathbb{P}$ nachází velké množství bodů odpovídajících terénu, pak lze náhodně vybírat trojice bodů, kde s vysokou pravděpodobností nalezneme trojici bodů odpovídajících terénu a tudíž nalezneme hledané řešení ve zlomku času potřebnému k otestování všech trojic bodů. Pro takové účely jsem použil algoritmus *RANSAC* [7], který je popsán v 3.2.3.

3.2.1 Kriteriální funkce

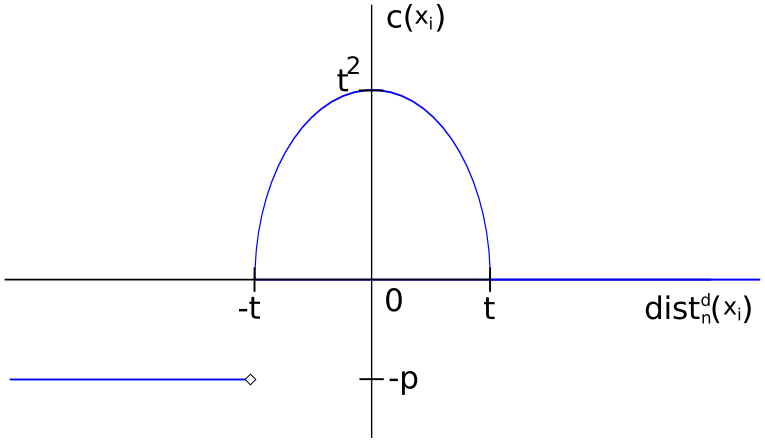
Základním kritériem pro určení kvality hypotézy lze považovat počet bodů ležících v rovině terénu. Dle výše uvedené definice roviny terénu uvedené v 3.1 platí, že v ideálním případě by se žádný bod neměl nacházet pod rovinou terénu. Ve skutečnosti množina bodů extrahovaná z disparitní mapy neodpovídá zcela realitě, jelikož je disparitní mapa odhadnuta s určitou nepřesností a proto odhadnutá pozice bodu v prostoru nemusí odpovídat skutečnosti a musíme ji brát s určitou nejistotou. Proto je tedy potřeba brát v potaz odchylky způsobené extrakcí bodů ze snímků a stanovit toleranční odchylku nad a pod rovinou terénu. Z tohoto důvodu také nelze požadovat nulový počet bodů nacházejících se pod rovinou a kriteriální funkci tedy nelze definovat pouze jako velikost průniku extrahovaných bodů s rovinou. Jednou z možných definic kriteriální funkce je počet bodů nacházející se v určitém pásu nad a pod rovinou.

Před definováním kriteriální funkce jsem zavedl vztahy

$$\begin{aligned} \mathbb{P} &= \{\mathbf{x} \,|\, \mathbf{x} \in \mathbb{R}^3\}\,, \\ \mathbb{I} &= \{\mathbf{x} \in \mathbb{P} \,|\, \mathrm{abs}(\mathrm{dist}_{\mathrm{d}}^{\mathbf{n}}(\mathbf{x})) \leq t\}\,, \\ \mathbb{O} &= \{\mathbf{x} \in \mathbb{P} \,|\, \mathrm{abs}(\mathrm{dist}_{\mathrm{d}}^{\mathbf{n}}(\mathbf{x})) > t\}\,, \\ \mathbb{U} &= \{\mathbf{x} \in \mathbb{O} \,|\, \mathrm{dist}_{\mathrm{d}}^{\mathbf{n}}(\mathbf{x}) < -t\}\,, \\ t &\in \mathbb{R}^+\,, \end{aligned} \tag{3.3}$$

kde $\mathbb{P}$ je množina odhadnutých bodů z disparitní mapy, $\mathbb{I}$–*inliers* ($\mathbb{O}$–*outliers*) odpovídají množinám bodů ležících v daném pásu (či mimo daný pás) roviny, jehož šířka je specifikována odchylkou t (*threshold*). Množina $\mathbb{U}$-*underliers* obsahuje body, které se nacházejí pod toleranční pásem roviny.

Nejjednodušší kriteriální funkce použitá pro stanovení kvality odhadu roviny terénu může být definována jako počet *inliers* $C = |\mathbb{I}|$. Aby kriteriální funkce přesněji vystihovala kvalitu nalezeného odhadu, je důležité, aby body v pásu nacházející se dále od roviny byly započítávány s menší váhou, jelikož je vhodné mít raději menší počet bodů blíže k rovině, než mnoho bodů ležících na okrajích tolerančního pásu roviny terénu. Pro započítávání váhy jednotlivých bodů v tolerančním páse jsem použil kvadratickou funkci. Dále aby se dle sekce 3.1 nalezená rovina nacházela v množině bodů co nejníže, zahrnul jsem tedy do kriteriální funkce penalizaci bodů ležících pod rovinou (respektive pod spodní hranicí tolerančního pásu).



Obrázek 3.1: Grafická reprezentace hodnoty kriteriální funkce pro vzdálenost jednoho bodu $\mathbf{x_i} \in \mathbb{P}$ od roviny

Kriteriální funkci pro jeden bod jsem definoval jako

$$C = \sum_{\mathbf{x_i} \in \mathbb{P}} c(\mathbf{x_i}) \,, \quad c(\mathbf{x_i}) = \begin{cases} t^2 - \mathrm{dist}_\mathrm{d}^\mathbf{n}(\mathbf{x_i})^2 \,, & \text{pokud } -t \leq \mathrm{dist}_\mathrm{d}^\mathbf{n}(\mathbf{x_i}) \leq t \,, \\ -p \,, & \text{pokud } \mathrm{dist}_\mathrm{d}^\mathbf{n}(\mathbf{x_i}) < -t \,, \\ 0 \,, & \text{jinak} \,, \end{cases} \tag{3.4}$$

kde p je penalizační konstanta jednoho bodu ležícího pod pásem roviny terénu.

Úhlové omezení Pro zvýšení pravděpodobnosti nalezení správné roviny terénu jsem zavedl úhlová omezení pro sklon roviny vůči osám x, y, d (popř. z). Vztahy pro výpočet sklonů jsou definovány jako

$$\begin{aligned} \gamma_x &= \mathrm{acos}(n_1) \,, \\ \gamma_y &= \mathrm{acos}(n_2) \,, \\ \gamma_z &= \mathrm{acos}(n_3) \,, \end{aligned} \tag{3.5}$$

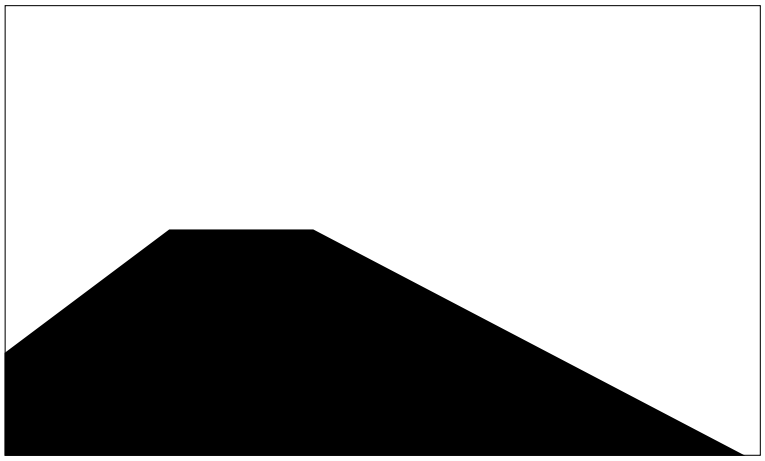
kde $\mathbf{n} = [n_1, n_2, n_3]^T$ je normálový vektor roviny terén a γ_x, γ_y, γ_z jsou velikosti úhlů mezi rovinou terénu a jednotlivými osami souřadného systému. Pro úhlová omezení tedy musí platit, že jsou splněny podmínky

$$\begin{aligned} \gamma_x &\in \langle \gamma_{min}^x, \gamma_{max}^x \rangle \,, \\ \gamma_y &\in \langle \gamma_{min}^y, \gamma_{max}^y \rangle \,, \\ \gamma_z &\in \langle \gamma_{min}^z, \gamma_{max}^z \rangle \,, \end{aligned} \tag{3.6}$$

kde intervaly $\langle \gamma_{min}^x, \gamma_{max}^x \rangle$, $\langle \gamma_{min}^y, \gamma_{max}^y \rangle$ a $\langle \gamma_{min}^z, \gamma_{max}^z \rangle$ určují povolený rozsah sklonu roviny pro danou osu. Pokud nejsou splněna všechna úhlová omezení, pak musí být hodnota kritéria roviny terénu rovna $-\infty$.

3.2.2 Použití masky vozovky

V případě hledání roviny terénu, kterou je vozovka, lze aplikovat na hustou disparitní mapu masku, která vymezuje část předpokládaného výskytu vozovky (viz. Obr. 3.2). Průnikem



Obrázek 3.2: Ukázka statické masky terénu

disparitní mapy a masky vznikne nová množina bodů P, ve které se nacházejí pouze body z oblasti vymezené maskou a hledání roviny terénu bude omezeno pouze na tuto podmnožinu. Použitím masky vozovky dojde k výraznému zmenšení množiny bodů, ale nedojde k odstranění optimálního řešení. Nejjednodušší použitelnou maskou je maska statická (viz. Obr. 3.2). Sofistikovanější alternativou by byla maska dynamická, která by se přizpůsobovala dle predikce pro následující snímek. Tato alternativa není v práci testována.

3.2.3 Odhad pomocí metody *RANSAC*

Množinou pozorování je množina bodů $\mathbb{P}$ získaná z disparitní mapy. Matematickým modelem, jehož parametry jsou odhadovány, je rovina. K získání roviny v prostoru $\mathbb{R}^3$ jsou potřeba tři body, kterými se rovina prokládá. Proto jsou v každé iteraci algoritmu *RANSAC* náhodně vybrány tři různé body z množiny extrahovaných bodů $\mathbb{P}$ a je spočítán normálový vektor $\mathbf{n}$ a posunutí v počátku d dle (3.1).

Takto odhadnutý model roviny je kontrolován pro splnění úhlového omezení (3.6) a v případě splnění podmínky je model vyhodnocen kriteriální funkcí dle (3.4), která určuje kvalitu získaného modelu. Poté je odhad uložen jako nejlepší možný nebo zamítnut v případě, že hodnota kriteriální funkce je nižší než hodnota předchozího nejlepšího odhad. V obou případech je iterační krok opakován a po překročení maximálního počtu iterací je algoritmus ukončen. Maximální počet iterací k je určen jako

$$k = \frac{\log(1-p)}{\log(1-w^n)}, \quad n = 3, \quad w = \frac{|\mathbb{I}|}{|\mathbb{P}|}, \tag{3.7}$$

kde n je počet náhodně vybíraných vzorků v jedné iteraci, w^n je pravděpodobnost, že všech n vybraných vzorků leží v rovině (*inliers*) a p odpovídá pravděpodobnosti, že získaný výsledek je přípustným řešením. Zvolili jsme hodnotu $p = 0.999$.

3.3 Stabilizace roviny pomocí Kalmanovy filtrace

Kalmanův filtr [5] je matematický nástroj umožňující získání nezašuměných dat ze zašuměných, nebo jinak znehodnocených dat. Tento algoritmus také umožňuje predikovat hodnotu v následujícím kroku a dá se tedy rozdělit na dvě části (predikční a korekční). Predikční část předpovídá, jak by měla data vypadat v následujícím kroku. Korekční část umožňuje pomocí naměřených dat opravit predikovaná data.

Za použití Kalmanovy filtrace predikujeme, kde by se měla v následujícím snímku (disparitní mapě) nacházet rovina terénu. Tato predikovaná rovina terénu je ohodnocena dle (3.4) a použita pro odhad roviny algoritmem *RANSAC* (sekce 3.2.3) jako zatím nejkvalitnější odhad roviny. Poté se algoritmus odhadu roviny terénu snaží najít kvalitnější reprezentaci roviny maximalizací kriteriální funkce. Po nalezení polohy roviny terénu je v následujícím kroku aplikován korekční krok Kalmanova filtru zajišťující korekci odhadu získaného metodou *RANSAC*. Tímto je zabráněno nepřiměřené změně polohy roviny terénu mezi dvěma snímky. Změna polohy odhadnuté roviny terénu v prostoru tedy bude plynulejší a zabrání se přijetí příliš vzdáleného špatného řešení, které může vzniknout chybami v disparitní mapě.

Predikce pro systém bez řízení je dána vztahy

$$\begin{aligned} \hat{\mathbf{x}}_{\mathbf{n+1}} &= \mathbf{A}\mathbf{x}_{\mathbf{n}}\,, \\ \hat{\mathbf{P}}_{n+1} &= \mathbf{A}\mathbf{P}_n\mathbf{A}^T + \mathbf{Q}\,, \end{aligned} \tag{3.8}$$

kde $\mathbf{x}$ charakterizuje vnitřní stav, $\mathbf{A}$ je matice přechodu do následujícího stavu. Dále je uvedena kovarianční matice $\mathbf{P}$ chyby odhadu a kovarianční matice $\mathbf{Q}$ šumu procesu. Pro korekci systému bez řízení platí

$$\begin{aligned} \mathbf{K} &= \hat{\mathbf{P}}_{n+1}\mathbf{C}^T(\mathbf{C}\hat{\mathbf{P}}_{n+1}\mathbf{C}^T + \mathbf{R})^{-1}\,, \\ \mathbf{x}_{\mathbf{n+1}} &= \hat{\mathbf{x}}_{\mathbf{n+1}} + \mathbf{K}(\mathbf{z} - \mathbf{C}\hat{\mathbf{x}}_{\mathbf{n+1}})\,, \\ \mathbf{P}_{n+1} &= (\mathbf{I} - \mathbf{K}\mathbf{C})\hat{\mathbf{P}}_{n+1}\,, \end{aligned} \tag{3.9}$$

kde $\mathbf{C}$ je matice převodu stavu na výstup, vektor $\mathbf{z}$ jsou naměřená data a matice $\mathbf{R}$ je kovarianční matice šumu měření. Pro získání výstupních dat z vnitřního stavu platí

$$\mathbf{y}_{\mathbf{n+1}} = \mathbf{C}\mathbf{x}_{\mathbf{n+1}}\,, \tag{3.10}$$

kde vektor $\mathbf{y}$ charakterizuje výstupní data. Konkrétní hodnoty jsou uvedeny v Tab. 3.1.

3.4 Odhad roviny terénu s omezující podmínkou na horizont

V některých případech lze požadovat nalezení roviny terénu, jež horizont bude rovnoběžný s horizontální hranou snímku[2]. V případě, že se vozidlo nebude pohybovat na zvláště deformované vozovce (např. závodní dráze), pak by měl být horizont vozovky (roviny terénu) v ideálním případě při zanedbání náklonu vozidla v zatáčkách rovnoběžný se spodní hranou snímku. Pro získání odhadu takovéto roviny s omezením na horizont jsou potřeba pouze dva body z prostoru $\mathbb{R}^3$. Z těchto dvou bodů je hodnota souřadnice x ignorována, jelikož

[2] předpokládáme stereopár umístěný vodorovně

$$\mathbf{x_t} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ \dot{n}_1 \\ \dot{n}_2 \\ \dot{n}_3 \end{bmatrix} \quad \mathbf{x_{t+1}} = \begin{bmatrix} n_1 + \dot{n}_1 \\ n_2 + \dot{n}_2 \\ n_3 + \dot{n}_3 \\ \dot{n}_1 \\ \dot{n}_2 \\ \dot{n}_3 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{P} = 10000 \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{R} = 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{Q} = 0.01 \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Tabulka 3.1: Konkrétní hodnoty matic Kalmanova filtru použité v práci

koresponduje se sklonem horizontu. Poté je možné proložit tyto dva body se se souřadnicemi $[y, z]^T$ přímkou charakterizovanou vektorem $\mathbf{m}$ a posunutím e, pro kterou platí vztah $[y, z]\mathbf{m} + e = 0$. Získání roviny v prostoru s fixní osou x z přímky v rovině lze provést rozšířením dimenze vektoru $\mathbf{m}$. Tudíž vektor roviny $\mathbf{n}$ bude odpovídat $[x, m_1, m_2]^T$ a posunutí $d = e$, kde $x = 0 => \mathbf{n} = [0, m_1, m_2]^T$. Konkrétní hodnoty jsou uvedeny v Tab. 3.2.

3.5 Vztah mezi modelem roviny terénu a disparitní mapou

Aby bylo možné vytvořit syntetickou disparitní mapu roviny terénu, kterou budeme potřebovat v kapitole 4, je nutno definovat vztah, který pro dané souřadnice $[x, y]^T$ syntetické disparitní mapy vrátí hodnotu disparity odpovídající rovině terénu. Mějme normálový vektor $\mathbf{n}$ a posunutí v počátku d. Potom mezi modelem roviny a disparitní mapou platí vztah

$$\begin{aligned} &n_1(x - x_c) + n_2(y - y_c) + n_3(d_{x,y}) + d = 0\,, \\ &\text{pro } 1 \leq x \leq w,\ 1 \leq y \leq h\,,\ d_{min} \leq d_{x,y} \leq d_{max}\,, \end{aligned} \tag{3.11}$$

kde $[x_c, y_c]^T$ jsou souřadnice optického středu hloubkové mapy, $d_{x,y}$ je disparita v bodě snímku $[x, y]^T$, d_{min} odpovídá minimální možné disparitě a d_{max} naopak maximální disparitě. Proměnná h značí výšku snímku a w šířku snímku v pixelech.

$$\mathbf{x_t} = \begin{bmatrix} n_1 \\ n_2 \\ \dot{n}_1 \\ \dot{n}_2 \end{bmatrix} \quad \mathbf{x_{t+1}} = \begin{bmatrix} n_1 + \dot{n}_1 \\ n_2 + \dot{n}_2 \\ \dot{n}_1 \\ \dot{n}_2 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{P} = 10000 \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{R} = 4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{Q} = 0.01 \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Tabulka 3.2: Hodnoty pro Kalmanovu filtraci pro rovinu terénu s omezující podmínkou na horizont

Získání disparity v bodě $[x, y]^T$ Pro vytvoření syntetické disparitní mapy musíme být schopni spočítat disparitu v každém bodě disparitní mapy. Proto definuji $\text{groundDisp}_{\text{d}}^{\mathbf{n}}(x, y)$ jako funkci pro získání disparity roviny terénu na pozici $[x, y]^T$ dle (3.11) jako

$$\begin{gathered} \text{groundDisp}_{\text{d}}^{\mathbf{n}}(x, y) = \max \left\{ d_{min} \ , \ -\frac{n_1(x-x_c)+n_2(y-y_c)+d}{n_3} \right\}, \\ \text{kde } 1 \leq x \leq w, \quad 1 \leq y \leq h \, . \end{gathered} \tag{3.12}$$

Výpočet řádku horizontu V kapitole 4 popisující metodu segmentace je potřeba znát řádek horizontu v_{hor} v syntetické mapě v daném sloupci x. Pro získání řádku horizontu v_{hor} platí, že hodnota disparity je rovna 0 proto tedy platí

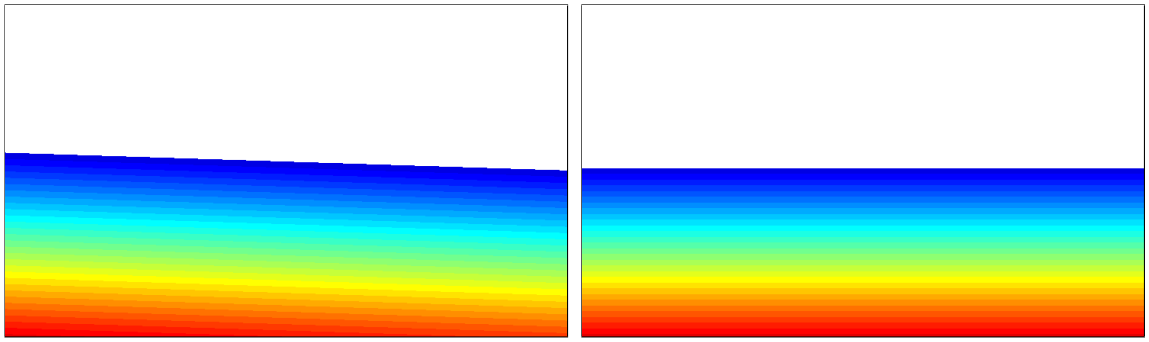
$$\begin{aligned} v_{hor}^x &= -\frac{n_1(x-x_c)+n_3*0+d}{n_2} + y_c \\ &= -\frac{n_1(x-x_c)+d}{n_2} + y_c \, . \end{aligned} \tag{3.13}$$

V případě roviny terénu s omezující podmínkou na horizont platí, že hodnota v_{hor} nezávisí na sloupci x a je v celém snímku konstantní.

Výpočet sklonu roviny v daném sloupci x Dále bude potřeba znát sklon α_x roviny terénu v daném sloupci x disparitní mapy, pro který platí vztah

$$\alpha_x = \frac{\text{groundDisp}_{\text{d}}^{\mathbf{n}}(x,h)}{h - v_{hor}} \, . \tag{3.14}$$

V případě roviny terénu s omezující podmínkou na horizont platí, že hodnota α nezávisí na sloupci x a je v celém snímku konstantní.



(a) Rovina terénu bez omezující podmínky na horizont

(b) Rovina terénu s omezující podmínkou na horizont

Obrázek 3.3: Syntetické disparitní mapy roviny terénu

Mapovací funkce Δ_d Funkce $\Delta_d(d_v, \Delta_Z)$ převádí odchylku Δ_Z na ose z v reálných jednotkách (m) na odchylku v (px) od dané disparity d_v.

$$\Delta_d(d_v, \Delta_Z) = \frac{bf}{\frac{bf}{d_v} + \Delta_Z} - d_v \,, \tag{3.15}$$

kde b je opět délka základny a f je ohnisková vzdálenost.

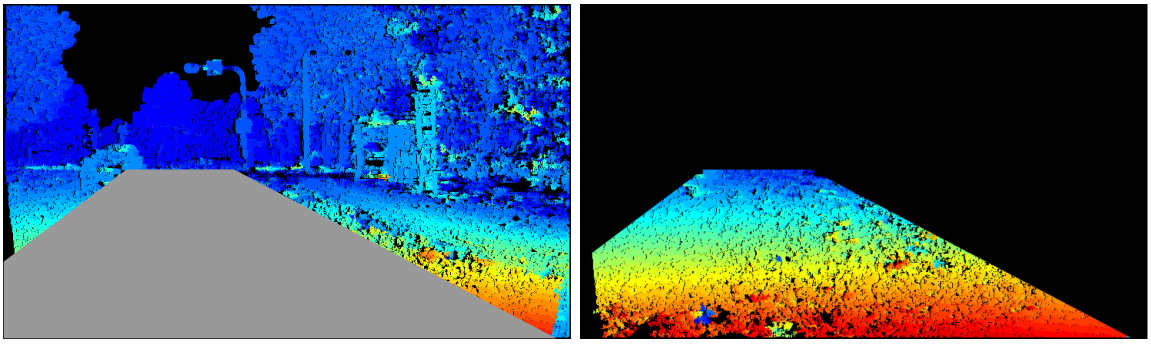
Vytvoření syntetické disparitní mapy roviny terénu Syntetická disparitní mapa roviny terénu odpovídá matici velikosti $w \times h$ s hodnotami určenými funkcí $\text{groundDisp}_{\text{d}}^{\mathbf{n}}(x, y)$, kde x a y jsou souřadnice pozice v matici (resp.mapě). Poté je disparitní mapa roviny terénu definována jako

$$\begin{bmatrix} \text{groundDisp}_{\text{d}}^{\mathbf{n}}(1,1) & \dots & \text{groundDisp}_{\text{d}}^{\mathbf{n}}(w,1) \\ \vdots & \ddots & \vdots \\ \text{groundDisp}_{\text{d}}^{\mathbf{n}}(1,h) & \dots & \text{groundDisp}_{\text{d}}^{\mathbf{n}}(w,h) \end{bmatrix} \tag{3.16}$$

3.6 Výsledky

Syntetická disparitní mapa s omezující podmínkou na horizont Obr. 3.3a charakterizuje syntetickou disparitní mapu bez omezující podmínky na horizont, kdežto Obr. 3.3b odpovídá syntetické mapě s omezující podmínkou na horizont, která zajišťuje, že horizont roviny terénu bude rovnoběžný s horizontální hranou disparitní mapy.

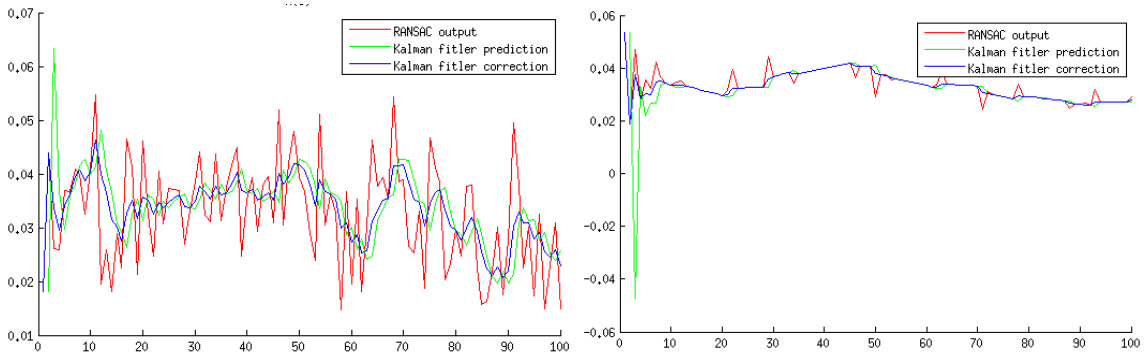
Použití masky roviny terénu Obr. 3.4a znázorňuje aplikaci masky vozovky (šedá barva) na disparitní mapu z Obr. 2.4a a Obr. 3.4b reprezetuje výslednou disparitní mapu vzniklou aplikováním masky, ve které se bude hledat rovina terénu. Aplikováním masky dojde k redukci počtu bodů, čímž dojde ke zvýšení pravděpodobnosti nalezení roviny terénu v menším počtu iterací. Dále vhodně navržená maska vozovky snižuje vliv okolního terénu mimo vozovku na hledanou rovinu terénu (v našem případě vozovky).



(a) Rovina terénu s vyznačenou oblastí masky (šedá část)

(b) Výsledná disparitní mapa pro hledání roviny terénu po aplikaci masky

Obrázek 3.4: Aplikace masky terénu na disparitní mapu



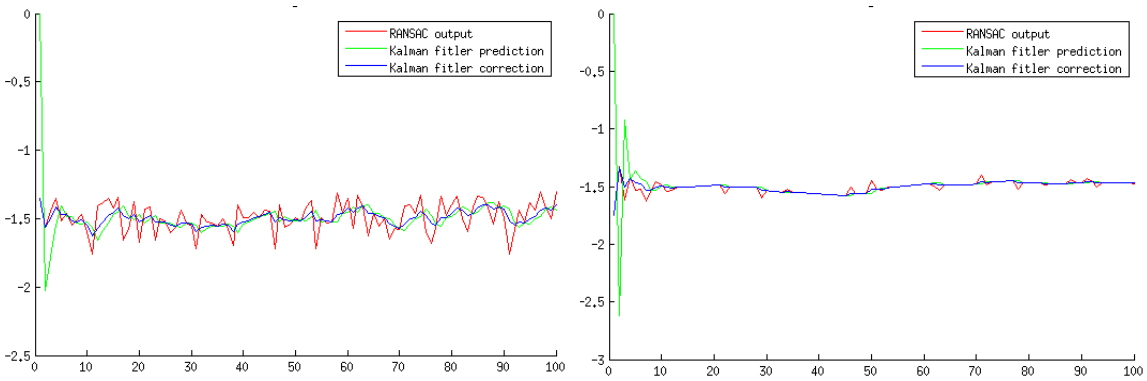
(a) Bez použití předikce

(b) S použitím předikce

Obrázek 3.5: 3.složka normálového vektoru n roviny terénu

Použití Kalmanova filtru Červená křivka v grafech Obr. 3.5 a Obr. 3.6 charakterizuje odhad parametrů roviny terénu pomocí algoritmu *RANSAC*. Zelená křivka charakterizuje predikci Kalmanovy filtrace a křivka modrá odpovídá výslednému odhadu modelu roviny terénu, který byl opraven korekcí získaného modelu metody *RANSAC*.

Obr. 3.5a a Obr. 3.6a odpovídají odhadu modelu roviny terénu bez použití predikce, kdežto Obr. 3.5b a Obr. 3.6b odpovídají hledání roviny terénu s použitím predikce jako zatím nejkvalitnější dosažený výsledek pro metodu *RANSAC*. Pokud algoritmus *RANSAC* nenalezne v dostatečném počtu iterací kvalitnější reprezentaci roviny terénu, pak použije predikci jako nejkvalitnější model roviny. Při využití Kalmanovy filtrace, jak je vidět na Obr. 3.5a a Obr. 3.6a, je změna polohy roviny v terénu v čase je plynulá a že použití Kalmanovy filtrace zamezuje skokové změně roviny. Dále se prokázalo, že použitím predikce pro metodu *RANSAC*, dojde k výraznému zrychlení hledání roviny terénu, protože pokud již existuje dostatečně kvalitní odhad roviny terénu, pak je algoritmus *RANSAC* ukončen po výrazně menším počtu iterací.



(a) Bez použití předikce (b) S použitím předikce

Obrázek 3.6: Posunutí d roviny terénu v počátku

Kapitola 4

Segmentace disparitní mapy

Tato kapitola popisuje modifikaci a implementaci metodu segmentace dle [6] za použití dynamického programování a zrychlení výpočtu pomocí vyhledávacích tabulek a paralelismu.

4.1 Charakteristika scény

Vstupními daty určenými k segmentaci jsou sekvence disparitních map pořízených z dopředu orientované stereokamery, která je umístěna na jedoucím vozidle. Každý snímek této sekvence zachycuje situaci silničního provozu v okamžiku pořízení. Snímky z dopravního prostředí obsahují charakteristické prvky jako je vozovka, dopravní značení (vertikální, horizontální - zejména semafory), ostatní účastníci provozu (dopravní prostředky, chodci, vozidla), přilehlý terén, horizont a budovy zasahující do okolního terénu. Pro segmentaci scény byla použita metoda [6].

4.2 *Stixel*ová reprezentace scény

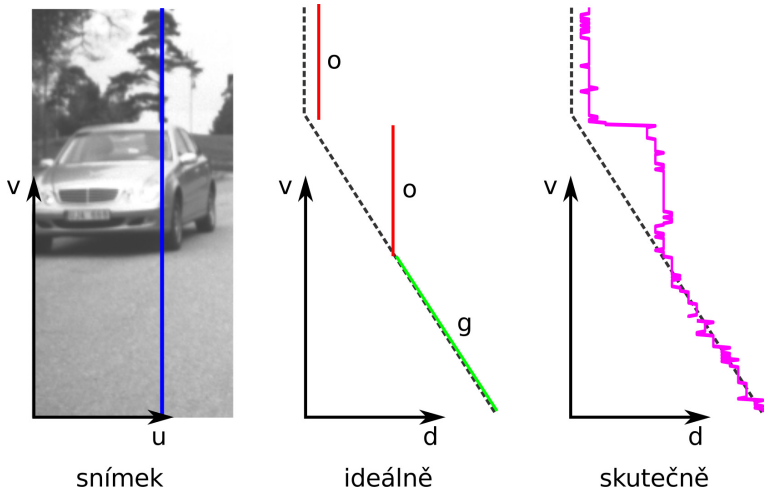
Mějme disparitní mapu $\mathbb{D}$ o rozměrech $w \times h$. Dle [1] rozdělme tuto disparitní mapu $\mathbb{D}$ na N_p vertikálních pruhů o šířce k pixelů, kde první pruh obsahuje prvních k sloupců disparitní mapy a každý následující sloupec obsahuje následujících k sloupců.

$$N_p = w/k\,,\ \text{za podmínky}\ k \in \mathbb{Z}\,, 1 \leq k \leq w\,, \tag{4.1}$$

kde N_p je počet pruhů, k je šířka pruhu a w je šířka disparitní mapy. Poté tedy dle [1] definujeme *Stixel* s_n (dle [1]) jako obdélníkovou část pruhu pro kterou dle [6] platí vztahy

$$\begin{array}{lll} L = \{L_u\}\,, & \text{kde} & 1 \leq u \leq N_p\,, \\ L_u = \{s_n\}\,, & \text{kde} & 1 \leq n \leq N_u \leq h\,, \\ s_n = \{v_n^b, v_n^t, c_n, f_n(v)\}\,, & \text{kde} & 1 \leq v_n^b \leq v_n^t \leq h\,, \quad c_n \in \mathbb{C}\,, \end{array} \tag{4.2}$$

kde s_n je n-tý *Stixel* v pruhu L_u, přičemž L_u je množina *Stixel*ů s_n v daném u-tém pruhu a N_u odpovídá celkovému počtu *Stixel*ů v pruhu u. Poté L je množina všech L_u a pro n-tý *Stixel* v pruhu L_u platí, že v_n^b je řádek začátku, v_n^t řádek konce, c_n je třída *Stixel*u



Obrázek 4.1: Grafická vizualizace ideální a skutečné hodnoty disparity objektu a roviny terénu.

korespondující s třídami klasifikace ze sekce 4.3. $f_n(v)$ odpovídá hodnotě funkce f na řádku v. Pro organizaci množiny *Stixel*ů L_u v u-tém pruhu dle [1] platí

$$1 = v_1^b \leq v_1^t \leq ... \leq v_{N_u}^b \leq v_{N_u}^t = h\,, \text{ kde} \quad v_{n-1}^b = v_t^n\,, \quad 1 < n \leq N_u. \tag{4.3}$$

Segmentace disparitní mapy pomocí *Stixel*ů dle [6] je rozdělení každého pruhu na n *Stixel*ů, kde pro zjednodušení a zrychlení výpočtu segmentace předpokládáme, že *Stixel*y v daném pruhu nejsou závislé na *Stixel*ech z ostatních pruhů, tudíž platí, že každá dvojice pruhů je nezávislá. Každý *Stixel* s_n (dle [6]) v daném pruhu u charakterizuje buď rovinu terénu g nebo planární objekt o. V rámci klasifikace celé disparitní mapy dochází díky využití *Stixel*ů k vysoké kompresi informací o segmentaci scény, protože každý pruh u je rozdělen na několik segmentů (*Stixel*ů) a není tedy třeba uchovávat informaci o klasifikaci každého pixelu. Takto reprezentovaná segmentace scény je vhodná pro další zpracování (např. vyhodnocení překážky).

4.3 Klasifikace *Stixel*u

Každý z prvků scény vyjmenovaných výše lze klasifikovat do dvou tříd. Nechť $\mathbb{C}$ je množina tříd, pak

$$\mathbb{C} = \{o, g\}\,, \tag{4.4}$$

kde o je planární objekt (*obstacle*) a g je rovina terénu (*ground*). Vozovku a přilehlý terén (odpovídající rovině terénu) lze klasifikovat jako g (terén) a zbylé prvky jako o (planární objekty). Dle [6] lze disparitu v každém sloupci disparitní mapy rozdělit na množinu lineárních funkcí – úseček (viz. Obr. 4.1) Pro klasifikaci *Stixel*u třídou o musí dle [6] platit, že každému řádku v celém segmentu vyhrazeném *Stixel*em s_n odpovídá stejná hodnota disparity a lze aproximovat lineární funkcí f_n^o. Poté pro $o \in \mathbb{C}$ platí

$$f_n^o(v) = \mu_n\,, \tag{4.5}$$

kde μ_n je průměrná hodnota disparity v daném *Stixel*u s_n a zároveň platí $v_n^b \leq v \leq v_n^t$. V případě klasifikace *Stixel*u s_n jako planární objekt o je uchovávána spolu s informacemi o umístění *Stixel*u s_n i informace o velikosti průměrné hodnoty disparity μ_n. Pro klasifikaci *Stixel*u třídou g musí dle [6] disparita na každém řádku segmentu *Stixel*u odpovídat disparitě roviny terénu a lze aproximovat lineární funkcí f_n^g. Poté pro $g \in \mathbb{C}$ platí vztah

$$f_n^g(v) = \alpha(v_{hor} - v)\,, \tag{4.6}$$

kde α je sklon roviny terénu, v_{hor} je řádek horizontu roviny terénu a pro v platí $v_n^b \leq v \leq v_n^t$.

Jelikož v této práci (na rozdíl od [6]) rovinu terénu odhadujeme v každém snímku, pak nemusí být v_{hor} a α ve všech snímcích a sloupcích konstantní.

4.4 Pravděpodobnostní model

Podrobný popis pravděpodobnostního modelu segmentace disparitní mapy pomocí *Stixel*ové reprezentace je popsán v [6]. Tato práce se dále bude zabývat implementací a modifikací metody segmentace za využití tohoto pravděpodobnostního modelu. Algoritmus segmentace bude implementován za použití dynamického programování v programovacím jazyce aplikace *Matlab*. Tato sekce obsahuje shrnutí vztahů z [6] a následně je doplněna o pravděpodobnost, která nebyla v [6] uvedena.

Hledání optimální segmentace disparitní mapy se realizuje maximalizací pravděpodobnostního ohodnocení všech pruhů disparitní mapy

$$L^* = \arg\max_{L \in \mathbb{L}} P(L|\mathbb{D})\,. \tag{4.7}$$

kde L je klasifikace pruhu a D je disparitní mapa. Použitím Bayesovy věty lze pravděpodobnost ohodnocení $P(L|\mathbb{D})$ převést

$$P(L|\mathbb{D}) \sim P(\mathbb{D}|L)P(L)\,. \tag{4.8}$$

Poté pro hledání klasifikace jednotlivých pruhů u platí

$$P(L|\mathbb{D}) \sim \prod_{u=1}^{N_p} P(D_u|L_u)P(L_u)\,, \tag{4.9}$$

kde N_p je počet *stixel*ů v pruhu u, D_u odpovídá disparitě pruhu u a L_u je segmentace pruhu u. Pro klasifikaci u-tého pruhu platí

$$\begin{aligned} P(D_u|L_u) &= \textstyle\prod_{n=1}^{N_u} \prod_{v=v_n^b}^{v_n^t} P(d_v|s_n, v)\,, \\ P(d_v|s_n, v) &= \begin{cases} P_D(d_v|s_n, v)\,(1 - p_\sharp^{c_n})\,, & \text{pokud } \exists(v) = 1\,, \\ p_\sharp^{c_n}\,, & \text{jinak} \end{cases}, \end{aligned} \tag{4.10}$$

kde c_n je třída klasifikace a $p_\sharp^{c_n}$ je pravděpodobnost, že daný pixel je díra. Pak platí

$$P_D(d_v|s_n, v) = \frac{p_{out}}{d_{max} - d_{min}} + \frac{\neg p_{out}}{A_{norm}} . \frac{1}{\sigma^{c_n}(d_v, v).\sqrt{2\pi}} . e^{\frac{1}{2}\left(\frac{d_v - f_n(v)}{\sigma^{c_n}(f_n, v)}\right)^2}\,, \tag{4.11}$$

$$A_{norm} = 0.5\Big(\mathrm{erf}\left(\frac{d_{max} - f_n(v)}{\sqrt{2}.\sigma^{c_n}(f_n, v)}\right) - \mathrm{erf}\left(\frac{d_{min} - f_n(v)}{\sqrt{2}.\sigma^{c_n}(f_n, v)}\right)\Big)\,, \tag{4.12}$$

kde $\sigma^{c_n}(f_n, v)$ v této práci odpovídá buď konstantě σ^o nebo σ^g. Dále platí

$$P(L_u) = P(s_1, \ldots, s_{N_u}) = P(s_1).\prod_{n=2}^{N_u} P(s_n|s_{n-1})\,. \tag{4.13}$$

Pro výpočet pravděpodobnostního ohodnocení prvního *Stixel*u v pruhu u platí

$$\begin{aligned} P(s_1) &= P(v_1^b, v_1^t, c_1, f_1(v)) \\ &= P(v_1^b).P(v_1^t).P(c_1|v_1^t).P(f_1(v)|c_1)\,, \end{aligned} \tag{4.14}$$

kde $P(v_1^b)$ zajišťuje začátek prvního *Stixel*u na prvním řádku pruhu u

$$P(v_1^b) = \begin{cases} 1\,, & \text{pokud } v_1^b = 1\,, \\ 0\,, & \text{jinak .} \end{cases} \tag{4.15}$$

Dále $P(v_1^t)$ určuje pravděpodobnost volby řádku konce daného *Stixel*u

$$P(v_1^t) = \tfrac{1}{h-1}\,. \tag{4.16}$$

Podmíněná pravděpodobnost $P(c_1|v_1^t)$ zabraňuje klasifikaci *Stixel*u nad horizontem jako rovinu terénu

$$P(c_1|v_1^t) = \begin{cases} 1\,, & \text{pokud } v_1^t > v_{hor}, c_1 = o\,, \\ 0\,, & \text{pokud } v_1^t > v_{hor}, c_1 = g\,, \\ 0.5\,, & \text{jinak .} \end{cases} \tag{4.17}$$

Podmíněný výraz $P(f_1(v)|c_1)$ modeluje pravděpodobnost aproximace disparity v *Stixel*u lineární funkcí planárního objektu či funkcí roviny terénu

$$P(f_1(v)|c_1) = \begin{cases} 1\,, & \text{pokud } c_1 = o\,, f_1(v) = \mu_1\,, \\ 1\,, & \text{pokud } c_1 = g\,, f_1(v) = \alpha(v_{hor} - v)\,, \\ 0\,, & \text{jinak .} \end{cases} \tag{4.18}$$

Pro výpočet pravděpodobnosti n-tého *Stixel*u v pruhu u platí

$$\begin{aligned} P(s_n|s_{n-1}) &= P(v_n^b|v_{n-1}^t).P(v_n^t|v_{n-1}^t).P(c_n|v_{n-1}^t, c_{n-1}). \\ &\quad P(f_n(v)|c_n, v_{n-1}^t, c_{n-1}, f_{n-1}(v))\,, \end{aligned} \tag{4.19}$$

kde výraz $P(v_n^b|v_{n-1}^t)$ zajišťuje začátek n-tého *Stixel*u s_n o řádek výše, než je řádek konce předchozího *Stixel*u s_{n-1}

$$P(v_n^b|v_{n-1}^t) = \begin{cases} 1\,, & \text{pokud } v_n^b = v_{n-1}^t + 1\,, \\ 0\,, & \text{jinak .} \end{cases} \tag{4.20}$$

Dále $P(v_n^t|v_{n-1}^t)$ určuje pravděpodobnost volby řádku konce *Stixel*u s_n v závislosti na řádku konce předchozího *Stixel*u s_{n-1}

$$P(v_n^t|v_{n-1}^t) = \begin{cases} 1/(h - v_{n-1}^t - 2)\,, & \text{pokud } v_n^t > v_{n-1}^t\,, \\ 0\,, & \text{jinak .} \end{cases} \tag{4.21}$$

c_n	c_{n-1}	Podmínka	P
o	o	$\mu_n > \mu_{n-1} + \Delta_d(\mu_{n-1}, \Delta_Z)$	$p_{ord}/(\mu_{n-1} - \Delta_d)$
o	o	$\mu_n \leq \mu_{n-1} - \Delta_d(\mu_{n-1}, \Delta_Z)$	$(1 - p_{ord})/(d_{max} - \mu_{n-1} - \Delta_d)$
o	o	$\lvert\mu_n - \mu_{n-1}\rvert < 2.\Delta_d(\mu_{n-1}, \Delta_Z)$	0
o	g	$\mu_n > f_{n-1}(v^t_{n-1}) + \epsilon$	$p_{grav}/(d_{max} - f_{n-1}(v^t_{n-1}) - \epsilon)$
o	g	$\mu_n < f_{n-1}(v^t_{n-1}) - \epsilon$	$p_{blg}/(f_{n-1}(v^t_{n-1}) - \epsilon)$
o	g	$\lvert\mu_n - f_{n-1}(v^t_{n-1})\rvert < 2\epsilon$	$(1 - p_{grav} - p_{blg})/2\epsilon$
g	o $\vee$ g	$f_n(v) = \alpha(v - v_{hor})$	1, jinak 0

Tabulka 4.1: Tabulka modelující pravděpodobnost $P(f_n(v)|c_n, v^t_{n-1}, c_{n-1}, f_{n-1}(v))$

c_n	c_{n-1}	v^t_n	P
o	o	$v^t_n > v_{hor}$	1
g	o	$v^t_n > v_{hor}$	0
o	g	$v^t_n > v_{hor}$	1
g	g	$v^t_n > v_{hor}$	0
o	o	$v^t_n \leq v_{hor}$	$1 - P_{og}$
g	o	$v^t_n \leq v_{hor}$	P_{og}
o	g	$v^t_n \leq v_{hor}$	1
g	g	$v^t_n \leq v_{hor}$	0

Tabulka 4.2: Pravděpodobnost $P(c_n|v^t_n, c_{n-1})$

Podmíněný výraz $P(f_n(v)|c_n, v^t_{n-1}, c_{n-1}, f_{n-1}(v))$ modeluje pravděpodobnost aproximace disparity v *Stixel*u s_n lineární funkcí planárního objektu či funkcí roviny terénu v závislosti na umístění a klasifikaci dvou po sobě jdoucích *Stixel*ů (s_n a s_{n-1}). Ohodnocení tohoto výrazu je uvedeno v Tab. 4.1.

Dle [6] jsem definoval funkci getDispProb pro výpočet pravděpodobnostní ohodnocení disparity v daném segmentu

$$\text{getDispProb}(v^b_n, v^t_n, fn_v) = \prod_{v=v^b_n}^{v^t_n} P(d_v|s_n, v)\,. \tag{4.22}$$

4.4.1 Navržení pravděpodobnosti

Tab. 4.2 charakterizuje pravděpodobnost $P(c_n|v^t_n, c_{n-1})$, která nebyla v práci [6] definována. Tabulka modeluje pravděpodobnost klasifikace *Stixel*u s_n v závislosti na řádku konce *Stixel*u s_n a třídě klasifikace c_{n-1} *Stixel*u s_{n-1}, kde platí, že nad horizontem terénu nesmí být *Stixel* klasifikován jako třída roviny terénu g. Dále platí, že výskyt roviny terénu za objektem o je méně pravděpodobný, než výskyt dalšího objektu o. Velikost této pravděpodobnosti výskytu roviny terénu g za objektem o modeluje konstanta P_{og}, pro kterou platí $P_{og} < 0.5$. Proměnná v_{hor} odpovídá řádku horizontu.

4.5 Implementace metody segmentace

4.5.1 Využití roviny terénu

Odhadnutá rovina terénu popsaná v kapitole 3 slouží v použité metodě segmentace pro vygenerování syntetické disparitní mapy terénu. Rozdíl disparit mezi disparitní mapou a syntetickou disparitní mapou terénu charakterizuje velikost odchylky disparitní mapy od odhadnutého modelu roviny terénu. Čím menší je odchylka od odhadnuté roviny, tím větší je pravděpodobnost, že se jedná o rovinu terénu.

4.5.2 Převod pravděpodobnosti na log-likelihood

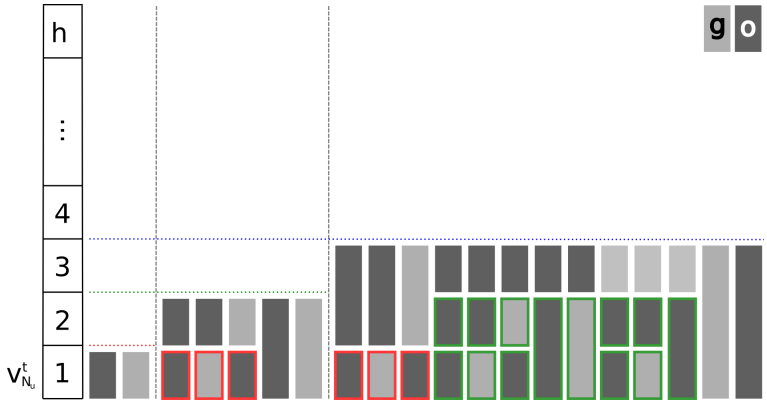
Aby bylo možné dále efektivněji pracovat s pravděpodobnostními ohodnoceními, pak je vhodné převést dané pravděpodobnosti na log-likelihood, který nám umožní sčítat jejich logaritmické hodnoty. Maximalizace pravděpodobnosti *Stixel*ové segmentace L^* dle definice 4.7 z [6] je maximalizace součinu pravděpodobnosti segmentace v daném pruhu u přes všechny pruhy

$$\begin{aligned} L^* &= \arg\max_{L\in\mathbb{L}} P(L|\mathbb{D})\,, \\ P(L|\mathbb{D}) &= P(\mathbb{D}|L).P(L) = \textstyle\prod_{u=1}^{N_p} P(D_u|L_u).P(L_u)\,. \end{aligned} \tag{4.23}$$

Pro maximalizaci pravděpodobnosti *Stixel*ové segmentace L_u pro pruh u platí dle [6] výraz

$$\begin{aligned} L_u^* &= \underset{L_u\in L,\, D_u\in\mathbb{D}}{\arg\max} P(D_u|L_u).P(L_u)\,, \\ &= \underset{L_u\in L,\, D_u\in\mathbb{D}}{\arg\max} \prod_{n=1}^{N_u}\prod_{v=v_n^b}^{v_n^t} P(d_v|s_n,v).P(s_1,\ldots,s_{N_u})\,, \\ &= \underset{L_u\in L,\, D_u\in\mathbb{D}}{\arg\max} \prod_{n=1}^{N_u}\prod_{v=v_n^b}^{v_n^t} P(d_v|s_n,v).P(s_1).\prod_{n=2}^{N_u} P(s_n|s_{n-1})\,, \\ &= \underset{L_u\in L,\, D_u\in\mathbb{D}}{\arg\max} \prod_{v=v_1^b}^{v_1^t} P(d_v|s_1,v).\prod_{n=2}^{N_u}\prod_{v=v_n^b}^{v_n^t} P(d_v|s_n,v).P(s_1).\prod_{n=2}^{N_u} P(s_n|s_{n-1})\,, \\ &= \underset{L_u\in L,\, D_u\in\mathbb{D}}{\arg\max} \left(P(s_1).\prod_{v=v_1^b}^{v_1^t} P(d_v|s_1,v)\right).\left(\prod_{n=2}^{N_u}\left(P(s_n|s_{n-1}).\prod_{v=v_n^b}^{v_n^t} P(d_v|s_n,v)\right)\right). \end{aligned} \tag{4.24}$$

Za účelem snadnější a přesnější práce s pravděpodobnostmi převedeme pravděpodobnostní model na logaritmickou věrohodnost (log-likelihood), čehož také využijeme při počítání



Obrázek 4.2: Grafická reprezentace všech možných způsobů uspořádání *Stixel*ů do výšky 3px. Červené ohraničení značí, že pro tyto segmenty se již v prvním kroku volala funkce pro vyhodnocení segmentace. Zelené ohraničení značí segmentace, které byly vyhodnoceny v druhém kroku.

výsledného ohodnocení segmentace.

$$\log\left(L_u\right) = \log\left(\left(P(s_1).\prod_{v=v_1^b}^{v_1^t} P(d_v|s_1,v)\right).\left(\prod_{n=2}^{N_u}\left(P(s_n|s_{n-1}).\prod_{v=v_n^b}^{v_n^t} P(d_v|s_n,v)\right)\right)\right) =$$

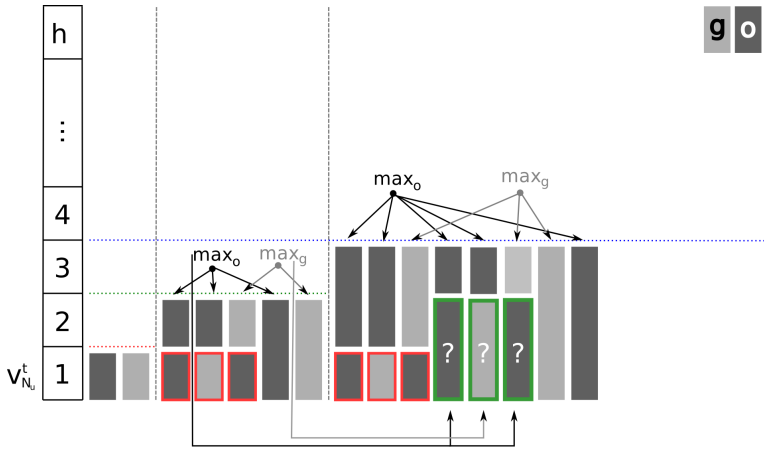
$$= \log\left(P(s_1).\prod_{v=v_1^b}^{v_1^t} P(d_v|s_1,v)\right) + \log\left(\prod_{n=2}^{N_u}\left(P(s_n|s_{n-1}).\prod_{v=v_n^b}^{v_n^t} P(d_v|s_n,v)\right)\right) =$$

$$= \log\left(P(s_1)\right) + \sum_{v=v_1^b}^{v_1^t}\log\left(P(d_v|s_1,v)\right) + \sum_{n=2}^{N_u}\left(\log\left(P(s_n|s_{n-1}) + \sum_{v=v_n^b}^{v_n^t}\log\left(P(d_v|s_n,v)\right)\right)\right) = \tag{4.25}$$

4.5.3 Nalezení optimální *Stixel*ové reprezentace sloupce

Obr. 4.2 ukazuje všechny možné způsoby uspořádání *Stixel*ů do výšky tří řádků. Již při výšce $v_{N_u}^t = 3$ řádků existuje 14 možností uspořádání, což odpovídá exponenciální výpočetní složitosti $2.2^{v_{N_u}^t}$. Pro disparitní mapu o výšce $h = 435$px to jest 2^{436} možností uspořádání a klasifikace *Stixel*ů do pruhu. Jak je patrné z Obr. 4.3, je možné při hledání segmentace pro řádek v využít segmentací spočítaných pro řádky $\{1,\ldots,(v-1)\}$. Při hledání segmentace pruhu o jeden řádek vyšší $(v+1)$ je dostačující nalézt uspořádání, které má nejvyšší možné ohodnocení pro každou ze tříd (o,g) v kombinaci s již hotovými uspořádáními pro jednotlivé řádky $\{1,\ldots,(v)\}$. Poté je pro disparitní mapu o výšce $h = 435$px potřeba pro nalezení optimální segmentace pouze 283836 volání funkce pro vyhodnocení segmentace oproti původních 2^{436} volání. Této vlastnosti rozdělení problému na podproblémy lze využít v dynamickém programování.

Dynamické programování Dynamické programování [3] nám umožňuje řešit složité úlohy za pomoci rozdělení komplexní úlohy na jednodušší podúlohy, kde hledání optima



Obrázek 4.3: Grafická reprezentace všech možných způsobů uspořádání *Stixel*ů do výšky 3px za použití již vypočítané segmentace o úroveň níže. Červené ohraničení značí nejlepší segmentaci do řádku 1 (včetně). Zelené ohraničení značí nejlepší segmentaci do řádku 2 (včetně).

Z	Do	Podmínka	Funkce ohodnocení hrany
$start$	o_n	$1 \leq n \leq h$	$f_{so}(n)$
$start$	g_n	$1 \leq n \leq h$	$f_{sg}(n)$
o_n	o_m	$1 \leq n < m \leq h$	$f_{oo}(n,m)$
o_n	g_m	$1 \leq n < m \leq h$	$f_{og}(n,m)$
g_n	o_m	$1 \leq n < m \leq h$	$f_{go}(n,m)$
$o_n \vee g_n$	$konec$	$n = h$	$f_k(n)$

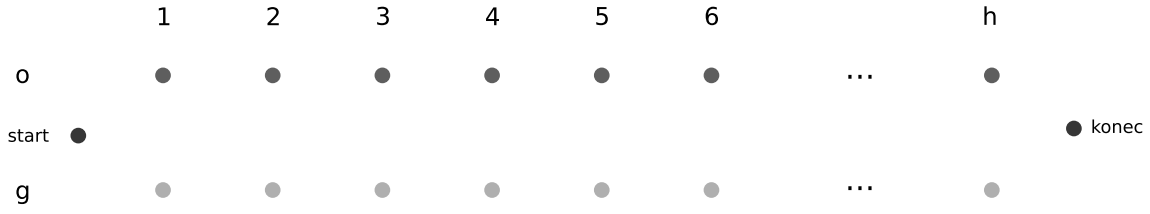
Tabulka 4.3: Tabulka definic přípustných hran

podúlohy vede k nalezení optima celé úlohy. V kontextu této metody segmentace to znamená, že pokud byla nalezena optimální segmentace části pruhu u do výšky v, pak lze efektivně nalézt optimální segmentaci pruhu do výšky $v+1$, kde po h iteracích nalezneme optimální segmentaci celého pruhu u. Hledání segmentace pomocí dynamického programování si lze představit jako hledání optimální cesty v orientovaném prostém grafu. Nechť je dána množina uzlů (viz. Obr. 4.4) $N = \{start, o_1, \ldots, o_h, g_1, \ldots, g_h, konec\}$. Cílem je nalézt cestu (viz. Obr. 4.5) z uzlu $start$ přes uzly o_n a g_n do uzlu $konec$ s nejvyšším možným ohodnocením. Definice přípustných hran je v Tab. 4.3. Uvedená definice přípustných hran zajistí, že z uzlu n do uzlu m existuje hrana pouze za podmínky $n < m$.

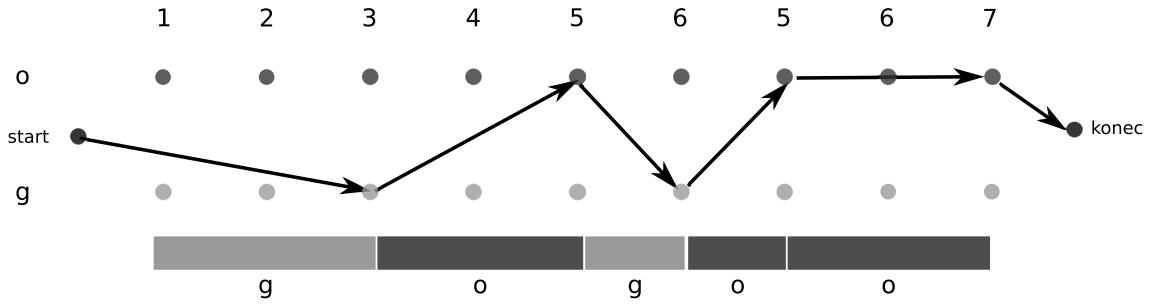
Ohodnocení hran mezi jednotlivými uzly V této sekci jsem za použití pravděpodobnostního modelu dle [6] odvodil ceny hran mezi jednotlivými uzly.

f_{so} - ohodnocení hrany z uzlu s do uzlu $o(v_1^t)$

$$f_{so}(v_1^t) = \log(\frac{1}{h-1} \cdot \begin{cases} 1\,, & \text{pokud } v_1^t > v_{hor}\,, \\ 0.5\,, & \text{jinak} \end{cases} \cdot \text{getDispProb}(1, v_1^t, \mu))\,. \tag{4.26}$$



Obrázek 4.4: Grafická reprezentace uzlů



Obrázek 4.5: Grafická reprezentace uzlů s nalezenou optimální cestou a segmentací

f_{sg} **- ohodnocení hrany z uzlu** s **do uzlu** $g(v_1^t)$

$$f_{sg}(v_1^t) = \log(\frac{1}{h-1}.\begin{cases}0\,, & \text{pokud } v_1^t > v_{hor}\,,\\ 0.5\,, & \text{jinak}\end{cases}.\,\text{getDispProb}(1, v_1^t, fn_v))\,. \tag{4.27}$$

$f_{oo}(v_{n-1}^t, v_n^t)$ **- ohodnocení hrany z uzlu** $o(v_{n-1}^t)$ **do uzlu** $o(v_{n-1}^t)$

$$\begin{aligned} f_{oo}(v_{n-1}^t, v_n^t) = &\log(\frac{1}{h-v_{n-1}^t-2}.\begin{cases}(1-P_{og}).\,, & \text{pokud } v_1^t \le v_{hor}\,,\\ 1\,, & \text{pokud } v_1^t > v_{hor}\end{cases}\\ &.\begin{cases}p_{ord}/(\mu_{n-1}-\Delta_d)\,, & \text{pokud } \mu_n > \mu_{n-1}+\Delta_d(\mu_{n-1},\Delta_Z)\,,\\ (1-p_{ord})/(d_{max}-\mu_{n-1}-\Delta_d)\,, & \text{pokud } \mu_n \le \mu_{n-1}-\Delta_d(\mu_{n-1},\Delta_Z)\,,\\ 0\,, & \text{pokud } |\mu_n-\mu_{n-1}| < 2.\Delta_d(\mu_{n-1},\Delta_Z)\end{cases}\\ &.\,\text{getDispProb}(v_{n-1}^t+1, v_n^t, \mu))\,. \end{aligned} \tag{4.28}$$

$f_{go}(v_{n-1}^t, v_n^t)$ **- ohodnocení hrany z uzlu** $g(v_{n-1}^t)$ **do uzlu** $o(v_{n-1}^t)$

$$\begin{aligned} f_{go}(v_{n-1}^t, v_n^t) = &\log(\frac{1}{h-v_{n-1}^t-2}.\begin{cases}1\,, & \text{pokud } v_n^t > v_{hor}\,,\\ 0.5\,, & \text{jinak}\end{cases}\\ &.\begin{cases}p_{grav}/(d_{max}-f_{n-1}(v_{n-1}^t)-\epsilon)\,, & \text{pokud } \mu_n > f_{n-1}(v_{n-1}^t)+\epsilon\,,\\ p_{blg}/(f_{n-1}(v_{n-1}^t)-\epsilon)\,, & \text{pokud } \mu_n < f_{n-1}(v_{n-1}^t)-\epsilon\,,\\ (1-p_{grav}-p_{blg})/2\epsilon\,, & \text{pokud } |\mu_n-f_{n-1}(v_{n-1}^t)| < 2\epsilon\end{cases}\\ &.\,\text{getDispProb}(v_{n-1}^t+1, v_n^t, \mu))\,. \end{aligned} \tag{4.29}$$

$f_{og}(v_{n-1}^t, v_n^t)$ **- ohodnocení hrany z uzlu** $o(v_{n-1}^t)$ **do uzlu** $g(v_{n-1}^t)$

$$\begin{aligned} f_{og}(v_{n-1}^t, v_n^t) = &\log(\frac{1}{h - v_{n-1}^t - 2} \cdot \begin{cases} 0\,, & \text{pokud } v_n^t > v_{hor}\,, \\ P_{og}\,, & \text{jinak} \end{cases} \\ &\cdot \begin{cases} 1\,, & \text{pokud } f_n(v) = \alpha(v - v_{hor})\,, \\ 0\,, & \text{jinak} \end{cases} \\ &\cdot \text{getDispProb}(v_{n-1}^t + 1, v_n^t, fn_v))\,. \end{aligned} \tag{4.30}$$

$f_{gg}(v_{n-1}^t, v_n^t)$ **- ohodnocení hrany z uzlu** $g(v_{n-1}^t)$ **do uzlu** $g(v_{n-1}^t)$

$$f_{gg}(v_{n-1}^t, v_n^t) = -\infty\,. \tag{4.31}$$

$f_k(v_n^t)$ **- ohodnocení hrany z libovolného uzlu do konce**

$$f_k(v_n^t) = \begin{cases} \log(1)\,, & \text{pokud } v_n^t = h\,, \\ -\infty\,, & \text{jinak} \end{cases}\,. \tag{4.32}$$

Hledání optimální cesty Každý uzel o_n a g_n uchovává informaci $o|g_{from}^n$ o svém předchůdci (předchozí uzel cesty), celkové ohodnocení cesty $o|g_{support}^n$ do daného uzlu a také hodnotu o_μ^n pro uzel o_n a hodnotu disparity g_g^n na řádku n v případě uzlu g_n. Uzel *konec* uchovává informaci $konec_{from}$ o svém předchůdci (o_h nebo g_h).

Princip hledání optimální cesty je ten, že n-tém kroku se hledá optimální cesta (segmentace) z uzlu *start* do uzlů o_n a g_n. Jelikož existují optimální cesty do všech předchozích uzlů ($o_1, \ldots, o_{n-1}$ a $g_1, \ldots, g_{n-1}$), pak je možné spočítat ohodnocení pro každou dostupnou cestu ze všech předchozích uzlů (včetně uzlu *start*) do uzlu o_n a vybrat cestu s nejvyšším ohodnocením. Stejně tak lze optimální cestu do uzlu g_n, pouze s tím rozdílem, že do uzlů g_i existují pouze cesty z uzlů o_i a uzlu *start* (zamítáme cestu z uzlu třídy g do uzlu třídy g).

Implementace hledání optimální cesty v pseudokódu popisuje příloha A. Segmentaci celého pruhu (cesty mezi uzly) lze získat zpětnou rekonstrukcí pomocí referencí $uzel from^n$ na předchozí uzel.

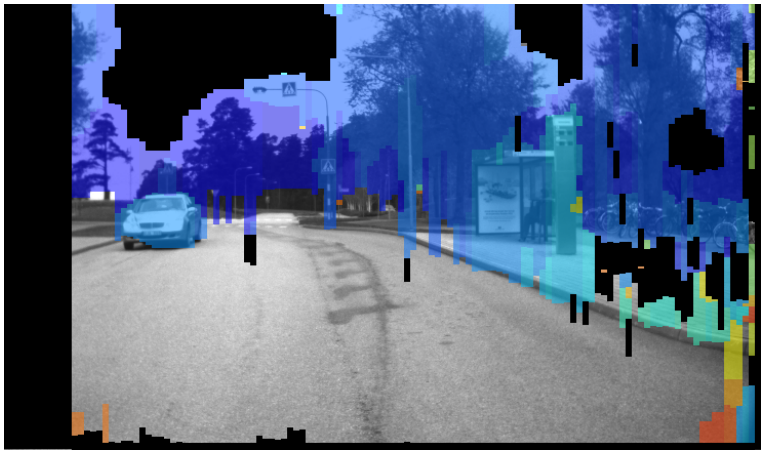
4.5.4 Zrychlení pomocí vyhledávacích tabulek

Pro zkrácení výpočetního času potřebného k segmentaci jsem využil možnost předpočítání hodnot opakovaně počítaných výrazů do tzv. vyhledávacích tabulek (look-up tables). Vyhledávací tabulka je datová struktura, která obsahuje výstupní hodnoty funkce f s n vstupními parametry. Kde parametry $p_i, i = 1..n$ indexují buňku vyhledávací tabulky. Vyhledávací tabulka se v počítačových programech nejčastěji realizuje pomocí n-rozměrného pole. Pokud nelze přímo indexovat buňky vyhledávací tabulky pomocí vstupních parametrů, lze vytvořit mapovací funkci, která převede dané vstupní parametry na index příslušné buňky ve vyhledávací tabulce.

Jako ukázku vyhledávací tabulky jsem zvolil předpočítání hodnot mapovací funkce Δ_d. Mapovací funkce Δ_d má dva vstupní parametry, jedním je konstanta Δ_Z, která charakterizuje odchylku v reálných jednotkách (metrech), druhým je proměnná d_v vůči které se hledá

	$+\Delta_Z$	$-\Delta_Z$
d_{min}	$\Delta_d(d_{min}, +\Delta_Z)$	$\Delta_d(d_{min}, -\Delta_Z)$
$d_{min}+1$	$\Delta_d(d_{min}+1, +\Delta_Z)$	$\Delta_d(d_{min}+1, -\Delta_Z)$
$\vdots$	$\vdots$	$\vdots$
d_{max}	$\Delta_d(d_{max}, +\Delta_Z)$	$\Delta_d(d_{max}, -\Delta_Z)$

Tabulka 4.4: Vyhledávací tabulka pro funkci Δ_d



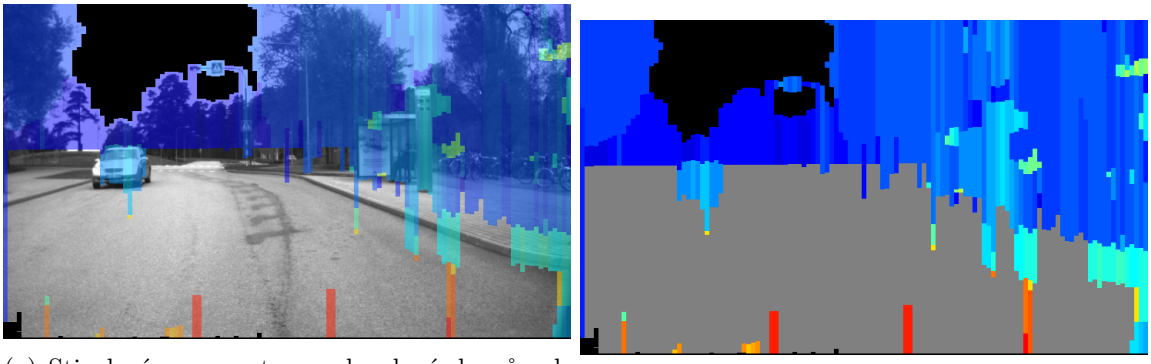
Obrázek 4.6: Ukázka možné segmentace

odchylka Δ_Z. Jelikož je potřeba jak odchylku kladnou $+\Delta_Z$, tak odchylku zápornou Δ_Z, pak bude mít vyhledávací Tab. 4.4 2 sloupce a n řádků, kde n je rozsah možných hodnot d_v, což odpovídá výrazu $n = d_{max} - d_{min} + 1$. Poté lze pro danou získat hodnotu odchylky pro danou disparitu d_v získáním hodnoty buňky $[d_v - d_{min} + 1, 1]$ pro $+\Delta_Z$ a $[d_v - d_{min} + 1, 2]$ pro $-\Delta_Z$ při indexování tabulky $[radek, sloupec]$ od 1. Dále jsou v implementaci zahrnuty vyhledávací tabulky pro pravděpodobnosti $P(v_n^t|v_{n-1}^t)$ a pro komplexní pravděpodobnost $P_D(d_v|s_n, v)$, kde předpokládáme konstantní hodnotu σ_o a σ_g v celém snímku. Tabulky zde nejsou uvedeny, jsou použity pouze v implementaci.

4.6 Výsledky

4.6.1 Výstup segmentačního algoritmu

Výstupem této metody segmentace je snímek (viz. Obr. 4.6) znázorňující segmentaci pomocí *Stixel*ů, kde odlišná barva určuje disparitu nacházející se v segmentu *Stixel*u. Černá barva charakterizuje díry v disparitních mapách. Dále je výstupem struktura obsahující parametry roviny terénu, což je normálový vektor a posunutí, a množina *Stixel*ů charakterizující planární objekty nalezené v disparitní mapě.



(a) Stixelová reprezentace vykreslená do původního snímku

(b) $\sigma_o = 3$, $\sigma_d = 3$

Obrázek 4.7: Segmentace disparitních map z *Growth*

4.6.2 Získání disparitní mapy použitím metody *Growth*

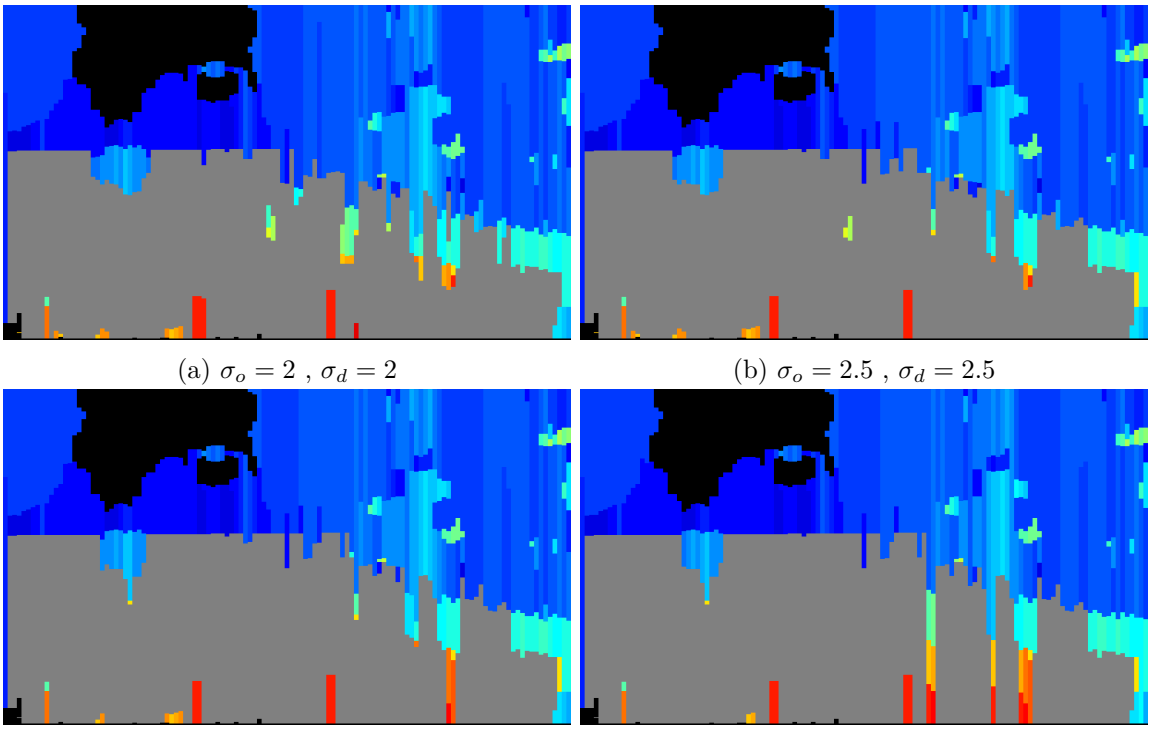
Tato sekce vyobrazuje sadu disparitních map získaných pomocí metody *Growth*. Obr. 4.8 zobrazuje několik výsledků segmentace s odlišnými parametry σ_o a σ_d. Obr. 4.7a charakterizuje *stixel*ovou segmentaci vykreslenou do původního snímku scény. Jak je patrné z níže uvedených obrázků segmentace, zvýšením hodnot σ_o a σ_d bude metoda segmentace tolerovat vetší odchylky a vytvářet tedy menší počet segmentů v daném sloupci.

4.6.3 Získání disparitní mapy pomocí vestavěné metody

Tato sekce obsahuje sadu disparitních map získaných pomocí vestavěné metody. Obr. 4.10 zobrazuje několik výsledků segmentace s odlišnými parametry σ_o a σ_d. Obr. 4.9a charakterizuje *stixel*ovou segmentaci vykreslenou do původního snímku scény. Jak je patrné z níže uvedených obrázků segmentace, zvýšením hodnot σ_o a σ_d bude metoda segmentace tolerovat vetší odchylky a vytvářet tedy menší počet segmentů v daném sloupci.

4.7 Diskuze

Možnosti zvýšení kvality segmentace Základním prvkem, ovlivňujícím výstupní kvalitu segmentace, je kvalita vstupní disparitní mapy. Čím je disparitní mapa hustěji a přesněji odhadnutá, tím je segmentace kvalitnější a více odpovídá realitě. Jelikož vstupními daty je sekvence, pak by bylo možné využít sekvenčního charakteru dat a vytvořit algoritmus pro získání disparitní mapy s použitím předchozí disparitní mapy. Tento algoritmus by zvyšoval kvalitu disparitních map a odstraňoval by tak alespoň z části šum, který se velmi často v jednotlivých snímcích vyskytuje, je pro každý snímek individuální, a také zvyšuje pravděpodobnost odhadnutí *Stixel*u jako neexistující objekt. Dále by bylo možné rozšířit pravděpodobnostní model metody segmentace o závislost jednotlivých *Stixel*ů na segmentaci přilehlých sloupců a na *Stixel*ech stejného sloupce ze segmentace předchozího snímku. Tak by bylo možné s určitou přesností redukovat výskyt *Stixel*ů odpovídajících šumu a ne skutečným objektům. Toto rozšíření by rozhodně zvýšilo kvalitu segmentace, bohužel na úkor výpočetní síly potřebné ke zpracování vstupních dat.



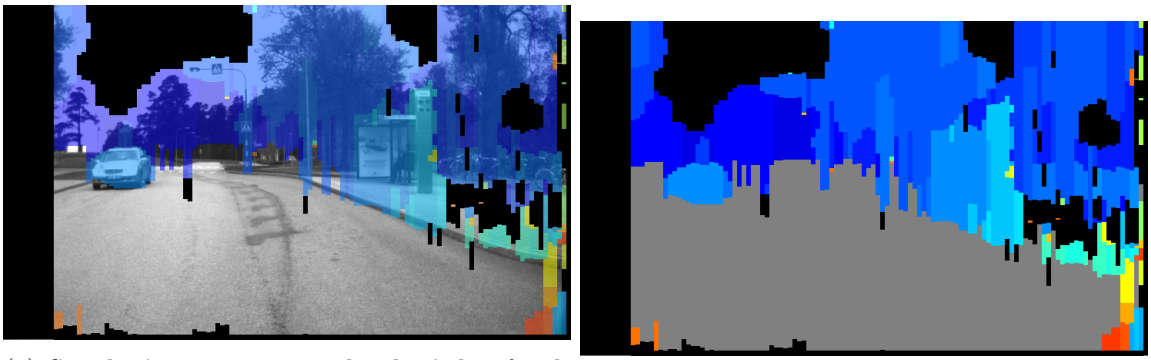
(a) $\sigma_o = 2$, $\sigma_d = 2$

(b) $\sigma_o = 2.5$, $\sigma_d = 2.5$

(c) $\sigma_o = 3$, $\sigma_d = 3$

(d) $\sigma_o = 3.5$, $\sigma_d = 3.5$

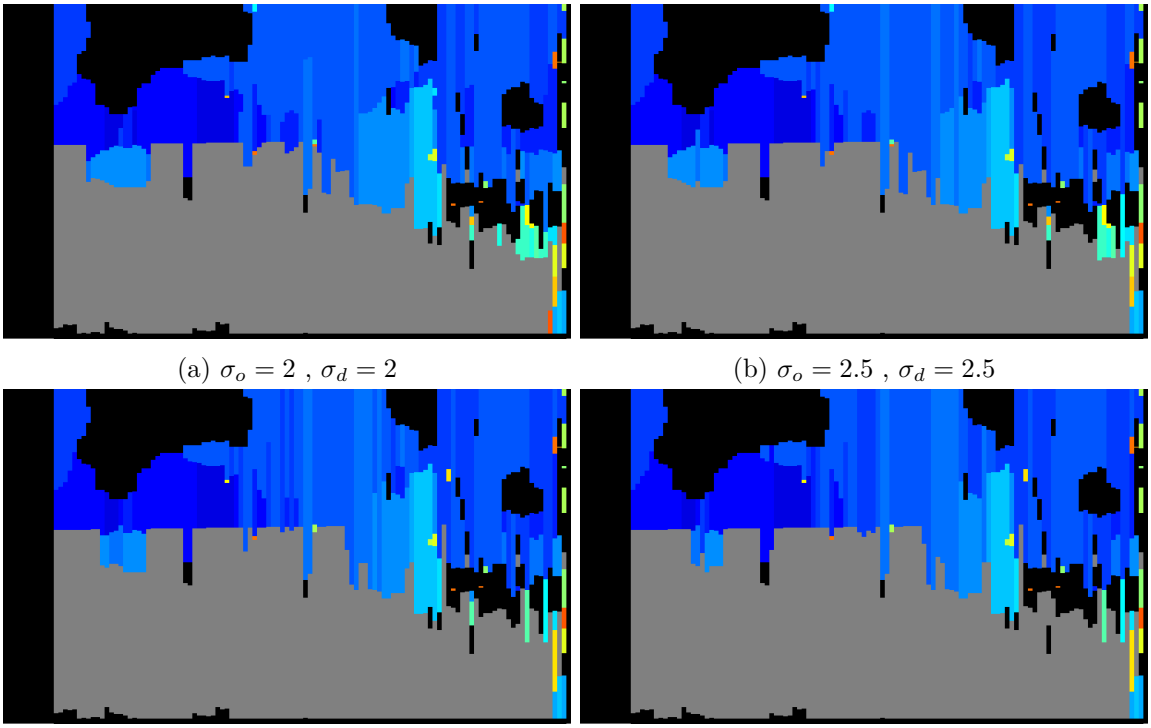
Obrázek 4.8: Segmentace pomocí metody *Growth* s různými σ_o a σ_d



(a) Stixelová reprezentace vykreslená do původního snímku

(b) $\sigma_o = 1$, $\sigma_d = 1$

Obrázek 4.9: Segmentace disparitních map získané vestavěnou metody



(a) $\sigma_o = 2$, $\sigma_d = 2$ (b) $\sigma_o = 2.5$, $\sigma_d = 2.5$

(c) $\sigma_o = 3$, $\sigma_d = 3$ (d) $\sigma_o = 3.5$, $\sigma_d = 3.5$

Obrázek 4.10: Segmentace pomocí vestavěné metody s různými σ_o a σ_d

Možnosti zrychlení procesu segmentace Pro implementaci této práce byl použit skriptovací jazyk (konkrétně *Matlab*), kde kód je interpretován a tudíž se nevyužívá plného potenciálu výpočetní síly procesoru. Použitý algoritmus je možné převést do nízkoúrovňového programovacího jazyka (např.C), který lze zkompilovat do strojového kódu procesoru, který lze efektivněji vyhodnocovat. Jelikož předpokládáme, že jednotlivé sloupce disparitní mapy nejsou závislé, pak lze vyhodnocovat několik sloupců ve stejném čase na více výpočetních jednotkách, čímž se dá za použití paralelismu dosáhnout výrazného zrychlení. Proto se nabízí možnost paralelního výpočtu segmentace ve vláknech vyhodnocovaných pomocí velkého množství procesorů grafické výpočetní jednotky GPU. Implementací algoritmu segmentace pomocí hardwarových prostředků je možné rapidně zvýšit rychlost výpočtu, kde by bylo možné provádět segmentaci v reálném čase.

Kapitola 5

Závěr

V této práci jsem stručně popsal problematiku získávání disparitních map a způsob extrahování bodů z disparitní mapy. Dále jsem navrhl a podrobně zdokumentoval metodu hledání roviny terénu pro kterou jsem definoval kriteriální funkci určující kvalitu nalezeného odhadu. Na takto získaný model roviny jsem použil filtraci, která zajistila stabilizaci roviny v čase a zabránila tak skokové změně polohy roviny. Následně jsem definoval funkci pro získání syntetické disparitní mapy z modelu roviny terénu. Samotná metoda segmentace je založena na práci [1], která poskytuje pravděpodobnostní model, který jsem rozšířil o neposkytnuté pravděpodobnosti. Tuto metodu jsem musel modifikovat, aby bylo možné použít odhadnutou rovinu terénu a implementovat ji. Implementaci samotné metody segmentace realizuji pomocí dynamického programování, které spolu s použitím vyhledávacích tabulek umožňuje výpočet segmentace v přijatelném čase. Jelikož jsem použil programovací jazyk *Matlab*, který je jazykem skriptovacím a jehož vyhodnocení trvá podstatně déle, než vyhodnocení nativního kódu dané architektury výpočetní jednotky CPU, tak se mi bohužel nepodařilo dosáhnout výpočtu v reálném čase. Aby bylo možné dosáhnout co nejvyšší kvality segmentace, musel jsem metodu odladit pro dané metody získávání disparitních map. Výsledky za použití odlišných metod získávání disparitních map, různých způsobů odhadu roviny terénu a různých parametrů metody segmentace jsem zdokumentoval a uvedl v jednotlivých kapitolách. Video obsahující vyhodnocení metody segmentace na sekvenci z reálného provozu je dostupné na přiloženém médiu. Na tomto médiu se také nachází celá implementace všech použitých metod.

Tuto práci je dále možné rozšířit o proces, který by na základě pravděpodobnostního modelu odstraňoval chyby v segmentaci způsobené četným šumem v disparitních mapách. Dalším možným pokračováním této práce je implementace metody segmentace do programovacího jazyka nižší úrovně s požadavkem na zpracování v reálném čase. Poté by bylo možné zde popsanou metodu použít například pro real-timovou detekci překážek v trajektorii vozidla a přizpůsobit tak chování vozidla dané situaci.

Výsledky segmentace zcela neodpovídají mým představám, určitě by bylo možné kvalitu metody segmentace ještě zvýšit, ale bohužel jeden semestr určený pro vypracování této práce je příliš málo. Každopádně práce na toto téma mě bavila a byla pro mě velkým přínosem, za což především děkuji svému vedoucímu práce, jehož cenné rady mi velmi pomohli.

Přílohy

Příloha A

Hledání optimální cesty

$o^1_{from} \leftarrow start$
$o^1_{support} \leftarrow f_{so}(1)$
$o^{mu}_1 \leftarrow mean_d(1,1)$

$g^1_{from} \leftarrow start$
$g^1_{support} \leftarrow f_{sg}(1)$
$g^1_d \leftarrow d(1)$

for all $\{i \in \mathbb{N} \mid 2 \leq i \leq h\}$ **do**

$o^i_{from} \leftarrow start$
$o^i_{support} \leftarrow f_{so}(i)$
$o^i_{mu} \leftarrow mean_d(1,i)$

$g^i_{from} \leftarrow start$
$g^i_{support} \leftarrow f_{sg}(i)$
$g^i_d \leftarrow d(i)$

for all $\{j \in \mathbb{N} \mid 1 \leq j < i\}$ **do**

$oo_support \leftarrow o^j_{support} + f_{oo}(j,i)$
$go_support \leftarrow g^j_{support} + f_{go}(j,i)$

if $oo_support > o^i_{support}$ **then**
$o^i_{from} \leftarrow o_j$
$o^i_{support} \leftarrow oo_support$
$o^i_{mu} \leftarrow mean_d(j+1,i)$
else if $go_support > o^i_{support}$ **then**
$o^i_{from} \leftarrow g_j$
$o^i_{support} \leftarrow go_support$
$o^i_{mu} \leftarrow mean_d(j+1,i)$
end if

$og_support \leftarrow o^j_{support} + f_{og}(j,i)$
if $og_support > g^i_{support}$ **then**

$g^i_{from} \leftarrow o_j$
$g^i_{support} \leftarrow og_support$
$g^i_d \leftarrow d(i)$
end if

end for

end for

if $support^o_h > support^g_h$ **then**
$konec_{from} \leftarrow o_h$
else
$konec_{from} \leftarrow g_h$
end if

Literatura

[1] BADINO, H. – FRANKE, U. – PFEIFFER, D. The Stixel World - A Compact Medium Level Representation of the 3D-World. In *Proc. 31st DAGM Symposium on Pattern Recognition*, 2009.

[2] DOBIÁŠ, M. – ŠÁRA, R. Real-Time Global Prediction for Temporally Stable Stereo. In *Proc. ICCV Workshop on Live Dense Reconstruction from Moving Cameras*, November 2011.

[3] *Dynamic programming* [online]. February 2013. Dostupné z: http://en.wikipedia.org/wiki/Dynamic_programming.

[4] HARTLEY, R. I. – ZISSERMAN, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[5] HAVLENA, V. – ŠTĚCHA, J. *Moderní teorie řízení*, 7.3 Kalmanuv filtr. Skriptum FEL, Vydavatelství ČVUT, 1996.

[6] PFEIFFER, D. – FRANKE, U. Towards a Global Optimal Multi-Layer Stixel Representation of Dense 3D Data. In *Proc. British Machine Vision Conference*, 2011.

[7] *RANSAC* [online]. February 2013. Dostupné z: http://en.wikipedia.org/wiki/RANSAC.

[8] RESEARCH, P. G. *PointGrey BumbleBee ®2* [online]. 2013. Dostupné z: http://ww2.ptgrey.com/stereo-vision/bumblebee-2.

[9] THE MATHWORKS, I. *Documentation Center: Computer Vision System Toolbox - disparity function* [online]. May 2013. Dostupné z: http://www.mathworks.com/help/vision/ref/disparity.html.

[10] THE MATHWORKS, I. *Documentation Center: Matlab Extension - mex function* [online]. May 2013. Dostupné z: http://www.mathworks.com/help/matlab/ref/mex.html.