

ŽILINSKÁ UNIVERZITA V ŽILINE

FAKULTA RIADENIA A INFORMATIKY

DIPLOMOVÁ PRÁCA

Bc. MILAN ONDRAŠOVIČ

Určovanie polohy statických objektov z pohyblivej kamery

Vedúci práce: Ing. Peter Tarábek, PhD.

Registračné číslo: 639/2018

Žilina, 2019

ŽILINSKÁ UNIVERZITA V ŽILINE

FAKULTA RIADENIA A INFORMATIKY

DIPLOMOVÁ PRÁCA

Informačné systémy – spracovanie dát

Bc. MILAN ONDRAŠOVIČ

Určovanie polohy statických objektov z pohyblivej kamery

Žilinská univerzita v Žiline

Fakulta riadenia a informatiky

Školiace pracovisko: Katedra matematických metód a operačnej analýzy (KKMOA)

Žilina 2019

ŽILINSKÁ UNIVERZITA V ŽILINE, FAKULTA RIADENIA A INFORMATIKY.

ZADANIE TÉMY DIPLOMOVEJ PRÁCE.

Študijný program : Informačné systémy

Zameranie: Spracovanie dát

Meno a priezvisko

Milan Ondrašovič

Osobné číslo

557543

Názov práce v slovenskom aj anglickom jazyku

Určovanie polohy statických objektov z pohyblivej kamery

Position estimation of static object from moving camera

Zadanie úlohy, ciele, pokyny pre vypracovanie

(Ak je málo miesta, použite opačnú stranu)

Cieľ diplomovej práce:

Cieľom práce je navrhnúť a implementovať metódu odhadu polohy statického objektu zachyteného pohyblivou kamerou (z auta). Na vstupe do systému bude sekvencia snímok zachytávajúca daný objekt, jeho detekcia vo forme ohraničujúceho obdĺžnika (anglicky bounding box) a trasovacie informácie prepájajúce výskyty objektu cez jednotlivé snímky. Zábery budú pochádzať z jednej kamery a budú k dispozícii aj údaje o polohe auta vo forme GPS. Navrhnuté riešenie bude otestované na reálnej úlohe určovania polohy dopravných značiek. Predpokladá sa tiež testovanie parametrov metódy na umelom datasete. Počas riešenia sa tiež preskúma možnosť využitia metódy odhadu polohy objektu na zvýšenie robustnosti trasovania.

Obsah:

1. Analýza súčasného stavu.
2. Návrh vhodnej metódy pre určovanie polohy statických objektov použitím jednej kamery.
3. Otestovanie metódy na reálnej úlohe určovania polohy dopravných značiek.
4. Preskúmanie možnosti využitia metódy odhadu polohy objektu na zvýšenie robustnosti trasovania.
5. Experimentálne vyhodnotenie v rôznych scenároch a záver.

Meno a pracovisko vedúceho DP: Ing. Peter Tarábek, PhD., KMMOA, ŽU

Meno a pracovisko tútora DP:

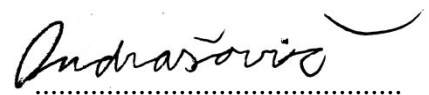
6.2.2019 
vedúci katedry
(dátum a podpis)

Zadanie zaregistrované dňa 26.10.2018 pod číslom 639/2018 podpis 

Čestné vyhlásenie

Vyhlasujem, že na práci „*Určovanie polohy statických objektov z pohyblivej kamery*“ som pracoval samostatne podľa pripomienok konzultanta a použil som len uvedenú literatúru.

Podpis autora práce:


.....

Pod'akovanie

Ďakujem svojmu konzulantovi za jeho vytrvalú spoluprácu, že mi zmenil môj pohľad na algoritmy a neposlednom rade za to, že ma zoznámil s oblasťou počítačového videnia.

ABSTRAKT V ŠTÁTNOM JAZYKU

ONDRAŠOVIČ, Milan: *Určovanie polohy statických objektov z pohyblivej kamery*. [Diplomová práca]. – Žilinská univerzita v Žiline. Fakulta riadenia a informatiky; Katedra matematických metód a operačnej analýzy (KMMOA). – Školiteľ/Vedúci: Ing. Peter Tarábek PhD. – Stupeň odbornej kvalifikácie: magister. – Mesto: FRI UNIZA, 2019. Počet strán: 80.

Zaoberáme sa problémom určovania pozície statického objektu z pohyblivej kamery. Konkrétne, určovaním pozície dopravného značenia z jednej kamery upevnenej na vozidle. Za týmto účelom boli vyvinuté dva matematické prístupy, pričom ako vstup predpokladáme anotáciu objektu vo forme ohraničujúceho obdĺžnika. Vytvorili sme syntetickú dátovú množinu, ktorá nám umožnila testovať metódy v kontrolovanom prostredí. Reálna dátová množina bola vytvorená nahrávkou trasy po meste. Výsledky ukázali, že v laboratórnych podmienkach je možné pozíciu odhadnúť s chybovosťou do 1 metra. Časté a presné GPS vzorkovanie je nevyhnutnou podmienkou, ktorá pri našich reálnych dátach nebola kompletne splnená, a toto rezultovalo v chybovosť počnúc menšou ako 1 meter až po v priemere 4 metre. Okrem iného sme predstavili aj návrh ako naše prístupy aplikovať pre vylepšenie trasovania objektov v zmysle kontrolného mechanizmu.

Kľúčové slová: statický objekt, pohyblivá kamera, lokalizácia, dátová množina

ABSTRAKT V CUDZOM JAZYKU

We are dealing with the problem of position estimation of static object from moving camera. Specifically, position estimation of traffic sign from single camera attached to a vehicle. To this end, two mathematical approaches have been developed, given the object annotation in the form of a bounding box is present on the input. We have created a synthetic data set, which enabled us to test the methods in a controlled environment. Real data set has been created as a recording in a town. Results have shown that in laboratory conditions the position can be estimated within the error of 1 meter. Frequent and accurate GPS sampling is a necessary condition, which has not been fully satisfied in our real data set, which resulted in an error range starting from below 1 meter up to 4 meters on average. Among other things, we have also suggested a potential application of our approaches for improving object tracking in a sense of a control mechanism.

Key words: static object, moving camera, localization, data set

Obsah

Zoznam obrázkov	9
Zoznam tabuliek	11
Zoznam skratiek	13
Úvod	14
1 Súčasný stav riešenej problematiky doma a v zahraničí	17
1.1 Základné pojmy a princípy danej oblasti	17
1.1.1 Proces tvorby fotografie	17
1.1.2 Azimut kamery	18
1.1.3 Zorné pole kamery.....	19
1.1.4 Hĺbková mapa.....	19
1.2 Popis rôznych iných metód	20
2 Metodika práce, metódy skúmania a návrh riešenia	27
2.1 Vytváranie syntetického <i>datasetu</i>	27
2.2 Vytváranie reálneho <i>datasetu</i>	32
2.3 Vzďialenosť od objektu pri priamočiarom pohybe.....	39
2.4 Spoločný základ pre lokalizáciu z dvoch snímok	42
2.4.1 Predpoklady.....	42
2.5 Určovanie polohy objektu na základe stretnutia priamok.....	44
2.6 Určovanie polohy objektu na základe triangulácie	47
2.7 Vylepšovanie odhadov jednotlivých metód	50
2.7.1 Globálny centroid pre odhad pozície.....	52
2.7.2 Zhlukovanie pre odstraňovanie šumu.....	52
2.8 Potenciál pre vylepšenie trasovania objektov	55
3 Výsledky práce a diskusia	57
3.1 Implementácia testovacej aplikácie.....	57
3.2 Metrika presnosti metód.....	57
3.3 Konverzia súradníc zemepisnej šírky a dĺžky do Kartézskeho súradnicového systému 58	
3.4 Testovanie, spracovanie a prezentovanie výsledkov.....	59
3.4.1 Výsledky testovania pre syntetické <i>datasety</i>	60
3.4.2 Sumárne štatistiky výsledných odhadov pre syntetické <i>datasety</i>	66
3.4.3 Výsledky testovania s umelým šumom pre syntetické <i>datasety</i>	67
3.4.4 Výsledky testovania pre reálny <i>dataset</i>	71
3.4.5 Testovanie chybovosti vzhľadom na vzdialenosť od objektu	72
Záver	75
Zoznam použitej literatúry	77

Prílohy	79
Príloha A: Obsah DVD	80

Zoznam obrázkov

Obrázok 1-1 Ilustrácia úbežníka.	18
Obrázok 1-2 Ukážka projekcie použitím <i>pinhole camera model</i> [12].	18
Obrázok 1-3 Ukážka projekcie použitím <i>pinhole camera model</i> [12].	18
Obrázok 1-4 Ilustrácia azimutu [14].	19
Obrázok 1-5 Ilustrácia zorného poľa kamery [18].	19
Obrázok 1-6 Príklad pôvodného obrázku (vľavo) a korešpondujúcej HM (vpravo) [10]. ..	20
Obrázok 1-7 Príklad stereogramu.	21
Obrázok 1-8 Stereo vizio: strom vykazuje väčšiu odchýlku ako dom, čiže je bližšie. [1, s. 2].	22
Obrázok 1-9 Geometrické závislosti stereo vizie [1, s. 2].	22
Obrázok 1-10 Triangulácia pomocou panorám GSW [3, s. 4].	22
Obrázok 1-11 Ukážka nerovnomerného rozloženia bodov pripadajúcich na 1 pixel [1, s. 3].	23
Obrázok 1-12 Ilustrácia štyroch rôznych pohľadov pri procese lokalizácie objektu. [4, s. 5].	24
Obrázok 1-13 Pri znalosti niekoľko strán objektu je pozícia vypočítateľná. [6, s. 3].	24
Obrázok 1-14 Ukážka hĺbkovej mapy pre vybrané obrázky [7, s. 7].	25
Obrázok 1-15 Lokalizácia chodca z troch pozícií kamery. [8, s. 2].	26
Obrázok 2-1 Zakrivená cesta, bez prevýšenia, rovný terén (<i>dataset flat</i>).	29
Obrázok 2-2 Zakrivená cesta, kopcovitý terén (<i>dataset hill</i>).	29
Obrázok 2-3 Zakrivená cesta, konštantné stúpanie (<i>dataset up</i>).	29
Obrázok 2-4 Rovná cesta, dopravné značenie umiestnené na ľavej strane (<i>dataset straight-left</i>).	29
Obrázok 2-5 Rovná cesta, dopravné značenie umiestnené na pravej strane (<i>dataset straight-right</i>).	30
Obrázok 2-6 Súradnicový systém SD.	30
Obrázok 2-7 Snímka 247 z <i>datasetu flat</i>	30
Obrázok 2-8 Ukážka RD – oblasť Žilina, blízko UNIZA.	33
Obrázok 2-9 Ukážka RD – oblasť Žilina, blízko UNIZA.	33
Obrázok 2-10 Ukážka procesu odhadovania súradníc konkrétneho miesta pomocou GM.	35
Obrázok 2-11 Snímka 112 z <i>datasetu zilina</i>	38

Obrázok 2-12 Ilustrácia odvodenia vzťahu pre výpočet vzdialenosti.	41
Obrázok 2-13 Ilustrácia výpočtu uhlu, pod ktorým je stred BB viditeľný.	44
Obrázok 2-14 Vizualizácia lokalizácie pomocou dvoch pomyslených priamok.	45
Obrázok 2-15 Situácia na prvej snímke.	47
Obrázok 2-16 Situácia na druhej snímke.	48
Obrázok 2-17 Konštrukcia pomysleného trojuholníka v danej scéne.	48
Obrázok 2-18 Trojuholník s uvedenými stranami a uhlami pre demonštráciu sínusovej vety.	49
3-1 Vizualizácia odhadov - <i>dataset: flat</i> , metóda: priamky.	60
3-2 Vizualizácia odhadov - <i>dataset: flat</i> , metóda: trojuholník.	61
3-3 Vizualizácia odhadov - <i>dataset: hill</i> , metóda: priamky.	61
3-4 Vizualizácia odhadov - <i>dataset: hill</i> , metóda: trojuholník.	62
3-5 Vizualizácia odhadov - <i>dataset: up</i> , metóda: priamky.	62
3-6 Vizualizácia odhadov - <i>dataset: up</i> , metóda: trojuholník.	63
3-7 Ukážka nerovnomerného rozloženia bodov pripadajúcich na 1 pixel [1, s. 3].	63
3-8 Vizualizácia odhadov - <i>dataset: straight-left</i> , metóda: priamky.	64
3-9 Vizualizácia odhadov - <i>dataset: straight-left</i> , metóda: trojuholník.	65
3-10 Vizualizácia odhadov - <i>dataset: straight-right</i> , metóda: priamky.	66
3-11 Vizualizácia odhadov - <i>dataset: straight-right</i> , metóda: trojuholník.	66
3-12 Vizualizácia chybovosti odhadov pozície vzhľadom na vzdialenosť objektu od kamery.	74

Zoznam tabuliek

Tabuľka 1 Popis parametrov SD.....	28
Tabuľka 2 Ukážka GPS záznamov.	32
Tabuľka 3 Parametre <i>datasetu zilina</i>	38
Tabuľka 4 Štatistika chybovosti (v metroch) jednotlivých odhadov pre <i>dataset flat</i>	60
Tabuľka 5 Štatistika chybovosti (v metroch) výsledného odhadu pre <i>dataset flat</i>	60
Tabuľka 6 Štatistika chybovosti (v metroch) jednotlivých odhadov pre <i>dataset hill</i>	61
Tabuľka 7 Štatistika chybovosti (v metroch) výsledného odhadu pre <i>dataset hill</i>	61
Tabuľka 8 Štatistika chybovosti (v metroch) jednotlivých odhadov pre <i>dataset up</i>	62
Tabuľka 9 Štatistika chybovosti (v metroch) výsledného odhadu pre <i>dataset up</i>	62
Tabuľka 10 Štatistika chybovosti (v metroch) jednotlivých odhadov pre <i>dataset straight-left</i>	64
Tabuľka 11 Štatistika chybovosti (v metroch) výsledného odhadu pre <i>dataset straight-left</i>	64
Tabuľka 12 Štatistika chybovosti (v metroch) jednotlivých odhadov pre <i>dataset straight-right</i>	65
Tabuľka 13 Štatistika chybovosti (v metroch) výsledného odhadu pre <i>dataset straight-right</i>	65
Tabuľka 14 Sumárna štatistika (v metroch) pre výsledné odhady pre syntetické <i>datasety</i>	67
Tabuľka 15 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre <i>dataset flat</i>	67
Tabuľka 16 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre <i>dataset flat</i>	68
Tabuľka 17 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre <i>dataset hill</i>	68
Tabuľka 18 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre <i>dataset hill</i>	68
Tabuľka 19 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre <i>dataset up</i>	69
Tabuľka 20 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre <i>dataset up</i>	69

Tabuľka 21 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre <i>dataset straight-left</i>	69
Tabuľka 22 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre <i>dataset straight-left</i>	70
Tabuľka 23 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre <i>dataset straight-right</i>	70
Tabuľka 24 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre <i>dataset straight-right</i>	70
Tabuľka 25 Štatistika chybovosti (v metroch) jednotlivých odhadov pre <i>dataset zilina (časť 1)</i>	71
Tabuľka 26 Štatistika chybovosti (v metroch) výsledného odhadu pre <i>dataset zilina (časť 1)</i>	71
Tabuľka 27 Štatistika chybovosti (v metroch) jednotlivých odhadov pre <i>dataset zilina (časť 2)</i>	71
Tabuľka 28 Štatistika chybovosti (v metroch) výsledného odhadu pre <i>dataset zilina (časť 2)</i>	71
Tabuľka 29 Štatistika chybovosti odhadov pozície vzhľadom na vzdialenosť objektu od kamery.	73

Zoznam skratiek

AP	Azimut pixelu
BB	Bounding Box
CSV	Comma Separated Values
CV	Computer Vision
FoV	Field of View
GM	Google Maps
GPS	Global Positioning System
GSW	Google Street View
GUI	Graphical User Interface
HM	Hĺbková mapa
JSON	JavaScript Object Notation
MVVM	Model-View-ViewModel
PSRNG	PSeudo-Random Number Generator
RD	Reálny <i>dataset</i>
RGB	Red-Green-Blue
SD	Syntetický <i>dataset</i>
WPF	Windows Presentation Foundation
XML	eXtensible Markup Language

Úvod

Vývoj v oblasti počítačového videnia (angl. *computer vision* (CV)) priniesol nemalé pokroky, najmä s rapídnyim znížením cien fotoaparátov či mobilných telefónov, ktoré aplikáciu podstatne uľahčujú. V dnešnej dobe je možné finančne nenáročnou cestou získať hardvér umožňujúci zaznamenávanie, uchovávanie ako aj dodatočné spracovávanie videa alebo fotografií. V každom z popísaných krokov sa počítačové videnie dokáže uplatniť, avšak predovšetkým v tom poslednom.

Mnohé úkony, ktoré sa ľuďom zdajú bežné, priam až samozrejmé, sú veľmi náročné pre počítače, a jedným z nich je vnímanie priestoru a hĺbky. Problém odhadovania pozície objektu zachyteného vo video zázname patrí medzi náročné problémy, ktorými sa oblasť počítačového videa zaoberá. Boli vyvinuté rôzne metódy, ktoré sa snažia tento problém riešiť s minimálnou chybovosťou, ktorá je pod vplyvom veľkého počtu faktorov.

Kameru možno chápať ako nástroj pomocou ktorého je dosiahnutá transformácia informácie z trojrozmerného priestoru do dvojrozmerného. Médium citlivé na svetlo umiestnené vo vnútri slúži na zachyt fotónov, ktoré sú od objektu záujmu odrazené. Takýmto médium môže byť elektronický senzor prípadne fotografický film. Elektronické kamery následne takúto informáciu prevedú do digitálnej formy, a táto je neskôr používateľovi najčastejšie prístupná vo forme obrázka.

Je potrebné si uvedomiť, že pri transformácií popísanej vyššie sa stráca informácia o hĺbke, resp. vzdialenosti. K danej transformácii teda neexistuje transformácia inverzná, ktorá by poskytla priestorové informácie spätne z obrázka. Projekcia z trojrozmerného priestoru do dvojrozmerného je teda stratová.

Strata priestorovej informácie je primárnou motiváciou nášho riešenia, nakoľko je potrebné vymyslieť a aplikovať metódu, ktorá dokáže túto informáciu aspoň čiastočne zrekonštruovať.

V našej práci sa zaoberáme odhadovaním pozície objektu za použitia video záznamu získaného kamerou. Nutnou podmienkou pre funkčnosť je, aby kamera polohu menila, pričom samotný objekt musí byť stacionárny. Rovnaké matematické závislosti by platili i v opačnom prípade, kedy by sa objekt pohyboval a kamera bola stacionárna.

Bez ujmy na všeobecnosti riešenia, náš problém sme špecifikovali ako lokalizáciu dopravného značenia z video záznamu získaného kamerou umiestnenou na automobile.

Prakticky je to úloha spracovať videosekvenciu obsahujúcu anotované dopravné značenia (do našej práce nespadá rozpoznávanie objektu), ktorých polohu je potrebné určiť vzhľadom na polohu kamery.

Predpokladom nášho riešenia je, aby dané vozidlo prešlo blízko značenia, ktorého polohu sa snažíme odhadnúť, a aby bola poskytnutá GPS poloha automobilu. Okrem týchto informácií je potrebné poznať aj aktuálne smerovanie vozidla. Táto informácia je vypočítaná na základe GPS pozícií.

Metódy uvedené v tejto práci sú výstupom úvah, ktorých predpoklady sme sa snažili minimalizovať. Výhodou takéhoto prístupu je schopnosť použiť ho v situáciách, kedy máme minimum prostriedkov, napríklad finančných, pomocou ktorých by sme mohli zaobstarat' kamery podporujúce stereo víziu. Naopak, nevýhodou je presnosť. Okrem iného, metódy takéhoto charakteru bývajú často implementačne jednoduchšie (relatívne k iným, robustným riešeniam), takže ďalším možným kritériom pre použitie môže byť aj časová náročnosť implementácie, ktorá je pri znalosti vyvođených vzťahov pomerne nízka.

Počas riešenia sa tiež preskúma možnosť využitia metódy odhadu polohy objektu na zvýšenie robustnosti trasovania. Vzniká tu potenciál vylepšiť proces trasovania nakoľko znalosť polohy aspoň s nejakou použiteľnou presnosťou by mohla pomôcť v prípade neistoty a predísť tak nesprávnej anotácii objektu v sekvenciách snímkov.

V nasledujúcej kapitole týkajúcej sa súčasného stavu riešenej problematiky popíšeme vedecké články, ktorými sme boli inšpirovaní, prípadne aj články, v ktorých sa zaoberajú príbuznými problémami a teda otvárajú možnosť pre prípadné rozšírenia v budúcnosti.

Následne podrobnejšie popíšeme ciele tejto práce, za ktorými prídu na rad samotné metódy, ktoré navrhujeme na riešenie. Okrem formalizovania matematických postupov budeme v neskorších kapitolách aj popisovať spôsob experimentovania a testovania konkrétnych metód ako na synteticky vytvorených dátach tak aj na dátach reálnych.

Disponujeme dátami vytvorenými v kontrolovanom prostredí, syntetickým *datasetom* (ďalej len SD), resp. dátovou množinou, avšak vždy budeme používať pojem *dataset*, ktoré nám umožňujú dosiahnuť maximálnu možnú presnosť. V takýchto podmienkach je možné metódu dôkladne testovať, kalibrovať, a následne vyvodit' závery či odhady, ako sa bude daná funkcionálna správať pri aplikácii na reálnom *datasete* (ďalej len RD).

Poskytujeme ukážky metódy na rôznych cestách (napríklad rovná, zakrivená a pod.), ako aj štatistické výsledky pre jednotlivé prípady. Metódu sme vždy dôkladne popísali spoločne s matematickými úvahami, na základe ktorých sú jednotlivé závislosti odvodené. Čitateľ by mal byť schopný pri dodržaní vyvodených matematických vzťahov získať identické výsledky, aké sme získali my.

Po testovaní za použitia SD je poskytnuté aj porovnanie s testami použitím RD vzhľadom na rôzne, nami definované, metriky.

Vízia do budúcnosti teda pozostáva z toho, že aj schopnosť samotného počítača použitím kamier pri vnímaní sveta priestorovo bude raz tak dobrá ako schopnosť človeka. V tejto práci sa k tejto vízií pokúsime o malý krok priblížiť.

1 Súčasný stav riešenej problematiky doma a v zahraničí

V tejto kapitole na úvod uvedieme niekoľko potrebných pojmov a princípov z oblasti počítačového videnia, ktoré považujeme za dôležité pre pochopenie jednotlivých metód, či už našich alebo iných autorov, ktoré budú rovnako v tejto kapitole uvedené.

Obmedzíme sa len na popis doménovo špecifických pojmov. Sekundárnym dôvodom pre zavedenie pojmov prípadne ich popisu je aj snaha vyhnúť sa nedorozumeniam, nakoľko množstvo literatúry je len v anglickom jazyku. Vyhradzuje si právo používať v texte anglické pojmy ak by ich preklad mal za následok možnú nekorektnú interpretáciu, avšak vždy bude snaha poskytnúť aj možný slovenský preklad (aspoň pri prvom použití).

1.1 Základné pojmy a princípy danej oblasti

1.1.1 Proces tvorby fotografie

Nutným predpokladom pre tvorbu fotografie je prítomnosť svetla, ktoré dopadá na objekt, ktorý je snaha zachytiť na fotografii. Svetlo po dopade na povrch objektu je odrazené smerom do optiky daného prístroja. Tam následne prechádza šošovkou až kým nedopadne na zobrazovacie médium, ktorým môže byť napríklad film fotoaparátu.

V modernej dobe sú využívané primárne digitálne fotoaparáty, ktoré používajú CMOS senzor na prevod svetelného signálu na digitálnu informáciu [11]. Svetelná informácia je zvyčajne popísaná tromi základnými farbami, červená (angl. *red*), zelená (angl. *green*) a modrá (angl. *blue*), na základe ktorých bol aj vytvorený tzv. RGB model. Existujú mnohé iné farebné modely, ale tie nie sú pre naše účely podstatné.

Jeden obrazový bod (angl. *pixel – picture element*) je teda tvorený trojicou čísel, ktoré reprezentujú intenzitu vyššie uvedenej trojice základných farieb. Obrázok sa teda chápe ako matica čísel, v ktorej prvok je vektor. Ak uvažujeme len jednu konkrétnu hodnotu vektora, potom hovoríme o matici, ktorá sa označuje ako kanál (angl. *channel*).

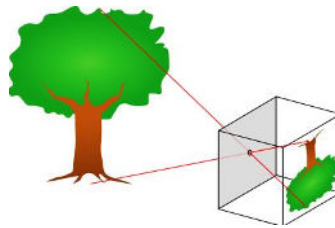
Formovanie obrazu z 3D na 2D priestor fotografie je matematicky vyjadriteľné ako projekcia. Existuje mnoho typov projekcií, avšak tá relevantná pre nás je tá, ktorou dokážeme popísať mechanizmus kamery.

Konkrétne sa jedná o perspektívnu projekciu, a je principiálne identická tej, ktorú používa ľudské oko. Jedným z príznakov takejto projekcie je tzv. úbežník (angl. *vanishing point*). Tento bod určuje priesečník dvoch rovnobežných priamok, ktoré po perspektívnej projekcii nie sú ďalej rovnobežné a teda sa v nejakom bode musia stretnúť.

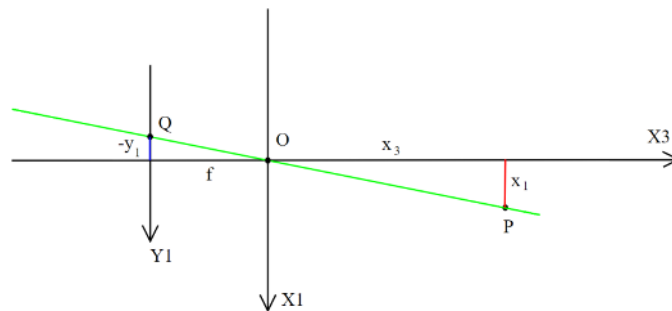


Obrázok 1-1 Ilustrácia úbežníka.

Geometricky je možné znázornenie pomocou modelu kamery, ktorý sa v angličtine označuje ako *pinhole camera model*. Pri projekcii sa stráca jedna dimenzia obsahujúca informáciu o hĺbke.



Obrázok 1-2 Ukážka projekcie použitím *pinhole camera model* [12].

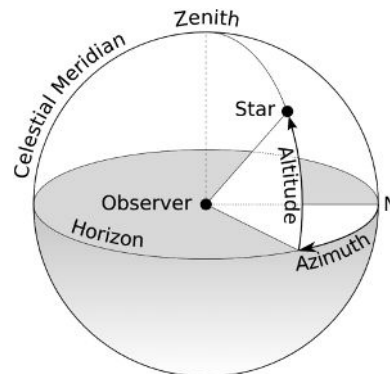


Obrázok 1-3 Ukážka projekcie použitím *pinhole camera model* [12].

1.1.2 Azimut kamery

Pod týmto pojmom budeme rozumieť uhlové vychýlenie kamery od nami vopred definovaného severu. Voči tejto polohe budeme schopný neskôr určiť skutočnú polohu objektu. Ako sever môže byť definovaný akýkoľvek smer, dôležité je len, aby bol pri výpočtoch korektne zahrnutý.

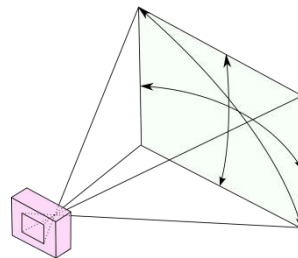
Tento pojem vychádza z astronómie, kde je definovaný ako uhol medzi južným bodom a priesečníkom výškovej kružnice hviezdy s horizontom; je to jedna z horizontálnych súradníc. [13].



Obrázok 1-4 Ilustrácia azimutu [14].

1.1.3 Zorné pole kamery

Pojem použitý pre horizontálny alebo vertikálny uhlový rozsah oblasti, ktorá je zachytená kamerou. Pokiaľ nebude v texte uvedené inak, v našom prípade implicitne predpokladáme len horizontálne uhlové zorné pole kamery.



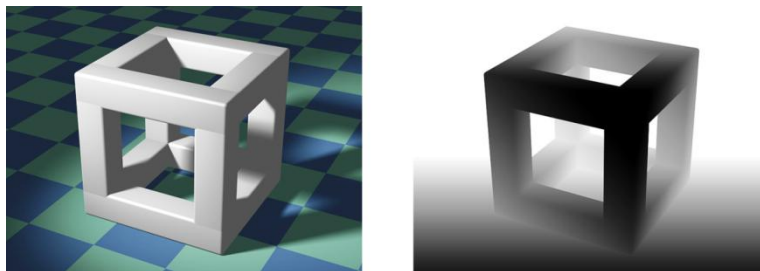
Obrázok 1-5 Ilustrácia zorného poľa kamery [18].

Zorné pole je časť obrazovej roviny objektívu ohraničená clonou alebo všeobecnejšie časť priestoru, ktorú možno optickým prístrojom zobrazit' [15].

Horizontálny uhlový rozsah kamery sa v anglickej literatúre označuje ako *horizontal field of view* (FoV), a vertikálny ako *vertical field of view* (FoV). Toto označenie budeme používať aj v našej práci.

1.1.4 Hĺbková mapa

Hĺbková mapa (angl. *depth map*) (ďalej len HM) predstavuje obrázok o identických rozmeroch ako je obrázok, ktorého HM počítame. Pre každý pixel daného obrázka je vypočítaná hĺbka a uložená na ekvivalentnej pozícii v rámci matice obrázka. Jedná sa teda o dvojrozmerný obraz, ktorý pre každý pixel pôvodného obrázku určuje vzdialenosť videného povrchu od pozorovateľa [10].



Obrázok 1-6 Príklad pôvodného obrázku (vľavo) a korešpondujúcej HM (vpravo) [10].

Takáto mapa je využívaná v rôznych algoritmoch, ktoré požadujú pre lokalizáciu globálnu hĺbkovú informáciu o celej scéne, ktorá je spracovávaná.

1.2 Popis rôznych iných metód

Vedecké publikácie obsahujú mnoho metód, ktoré sú použiteľné pre riešenie problému odhadu pozície objektu. Existujú konkrétne metódy pre monokulárne či binokulárne kamery pohybujúce sa v priestore, prípadne konfigurácie s viacerými kamerami. Umiestnenie kamery je rovnako dôležitým faktorom a býva často špecifické pre konkrétne riešenie. V našom prípade odhadu pozície dopravného značenia by umiestnenie v strede automobilu, prípadne na pravej strane bolo ideálne, nakoľko sa dopravné značenie vyskytuje na pravej strane vozovky.

Snaha využívať len vizuálnu informáciu pre lokalizáciu je cesta relatívne jednoduchá, avšak pri spracovávaní problematická. Rolu nehrá len pohyb vozidla a zmena obrazu spôsobená fyzickým pohybom kamery, ale nezanedbateľný vplyv na kvalitu spracovania obrazu má aj osvetlenie či rôzne druhy šumu, ktoré môžu výsledky dramaticky ovplyvniť. Príkladom môže byť nedostatočné osvetlenie v čase hmly alebo dažďa. Naša úloha je podstatne sťažená aj potenciálnymi nepresnými hodnotami v GPS súradniciach. Predpokladom teda je, že získané údaje nie sú dokonalé.

Ak abstrahujeme od samotnej chyby videa po optickej stránke, negatívny vplyv môže mať aj nekorektná detekcia objektu záujmu, na ktorú vplyv nemáme, nakoľko predpokladom na vstupe je anotovaná videosekvencia.

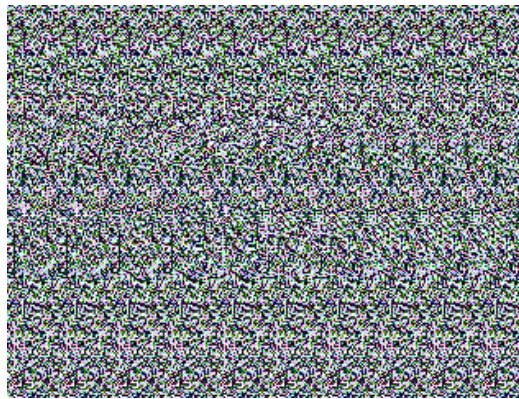
Jednoduchším problémom odhadovania pozície objektu vzhľadom na kameru je odhadovanie vzdialenosti objektu od kamery. Pre naše riešenie nie je potrebná len vzdialenosť objektu, ale aj jeho absolútna/relatívna poloha vzhľadom na polohu kamery v čase získania daného snímku. Nakoľko primárnym cieľom našej práce je určiť polohu objektu v 2D rovine tak elevácia (prevýšenie) nie je potrebná.

Okrem iného sa zmienime aj o sofistikovaných metódach, ktorými je možné sa inšpirovať do budúcnosti pri vývoji ešte robustnejšej metódy, nakoľko táto problematika je otvorená.

Autori článku [1] použili stereo víziu pre odhadovanie polohy objektu. Tento postup spočíva v použití dvoch kamier laterálne vychýlených, kde zmena polohy objektu relatívne k jeho pozadiu pri jednotlivých snímkach je postačujúca informácia pre výpočet vzdialenosti [1, s. 2].

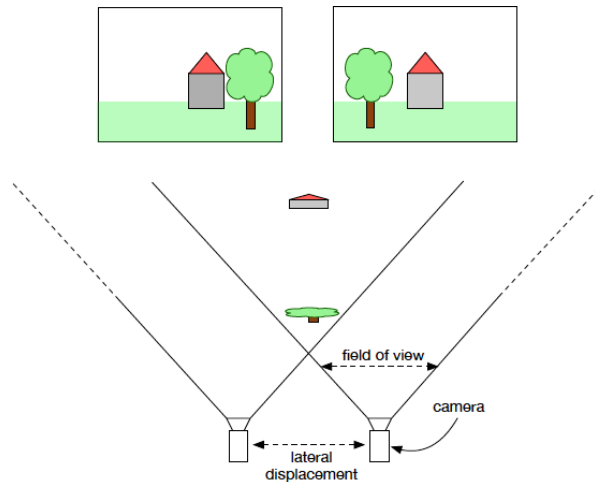
Na základe tejto informácie sú aj ľudia schopní odhadovať vzdialenosti k objektom pomocou oboch očí. Je teda možné tvrdiť, že binokulárne, stereoskopické videnie je závislé na schopnosti nervového systému registrovať prítomnosť malých rozdielov v pozícií objektov vzhľadom na ľavú a pravú očnú sieťnicu [2, s. 2].

V niekoľkých výskumoch bolo zistené, že človek vníma hĺbku rozdielom v pohľade z ľavého a pravého oka. Jedným zo spôsobov demonštrácie tohto fenoménu je tzv. *Random Dot Stereogram* [9], pomocou ktorého sa účastník pozerá na obrázok obsahujúci napríklad bodky náhodne rozmiestnené a dostáva pocit, že sa obrázok hýbe, prípadne sa mu zdá, že je trojrozmerný.



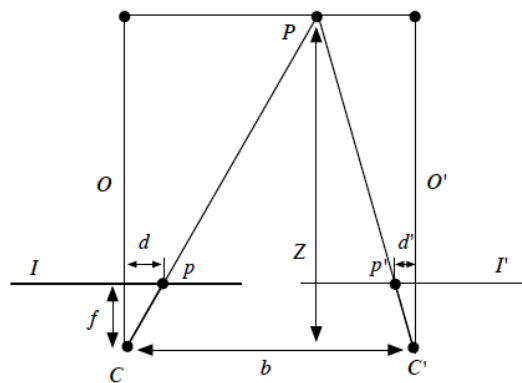
Obrázok 1-7 Príklad stereogramu.

Tento prístup sme ale nevyužili v jeho pravej forme, ale snažili sme sa pracovať len s použitím jednej kamery. Naša snaha sa dá interpretovať ako pokus nahradiť priame riešenie stereo víziou riešením jednoduchším, kde pomocou jednej kamery získame čo najviac informácií z obrazu, aby bolo možné aplikovať princípy stereo vízie, aj keď za cenu menej kvalitného a presného riešenia.



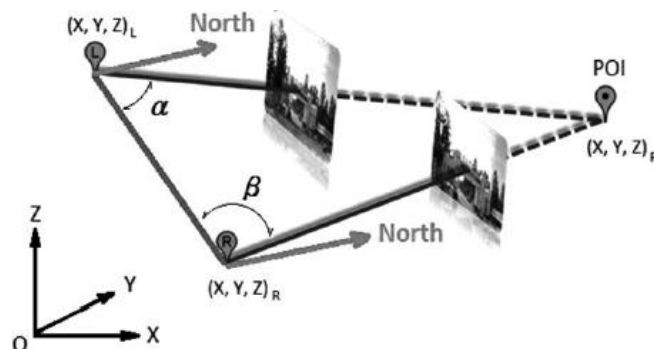
Obrázok 1-8 Stereo vzia: strom vykazuje väčšiu odchýlku ako dom, čiže je bližšie. [1, s. 2]

Na obrázku 1-8 je ilustrovaná stereo vzia v praxi. Je nutné poznať laterálnu odchýlku oboch kamier (tých, ktoré sú umiestnené priamo „vedľa seba“), ako aj ohniskovú vzdialenosť oboch kamier, pričom sa odporúča, aby kamery boli identické [1, s. 2].



Obrázok 1-9 Geometrické závislosti stereo vzie [1, s. 2].

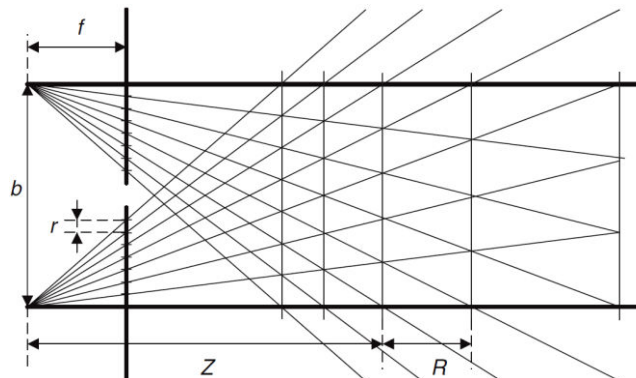
Prístup podobný nášmu využili autori článku [3]. Ich cieľom bolo odhadovať vzdialenosť k ľubovoľným objektom pomocou dvoch panoramatických záberov z *Google Street View* (ďalej len GSW). Ich riešením bolo použitie triangulácie, ktoré vo veľkej miere reflektuje aj závislosti, ktoré sme využili i my.



Obrázok 1-10 Triangulácia pomocou panorám GSW [3, s. 4].

Pri výpočte pozície bola využitá jednoduchá závislosť, ktorá platí pre trojuholníky známa ako sínusová veta. Táto veta umožňuje pomocou známej jednej strany a dvoch uhlov vypočítať dĺžky ostatných strán.

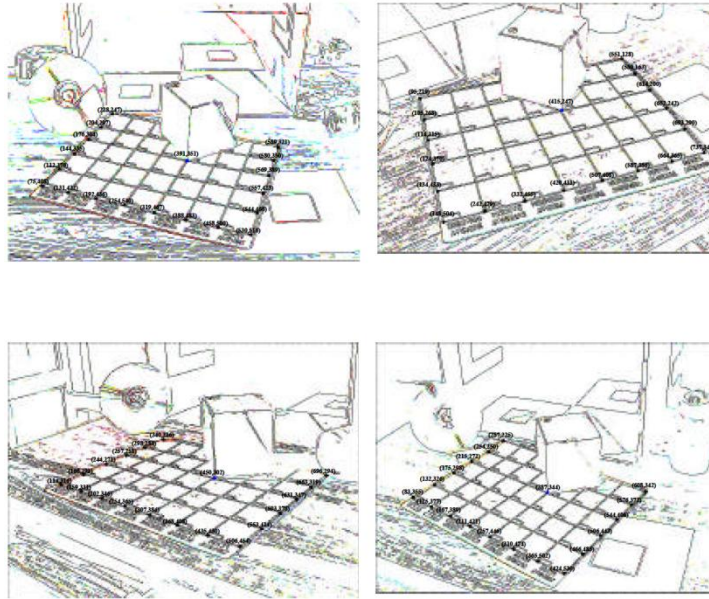
Táto metóda je v konečnom dôsledku matematicky jednoduchá, primárnou nevýhodou je len samotná nepresnosť GPS prípadne granularita obrázku, pretože malé rozlíšenie predstavuje veľké množstvo potenciálnych bodov v skutočnom priestore, ktoré pripadajú na jeden pixel [1, s. 3].



Obrázok 1-11 Ukážka nerovnomerného rozloženia bodov pripadajúcich na 1 pixel [1, s. 3]

Výsledky autorov vykazovali očakávané hodnoty. Chyba v prípade súradnice X bola približne 0,3 metra, v prípade súradnice Y približne 1,8 metra, avšak v prípade Z chyba bolo až okolo 5 metrov. Tento výsledok je očakávaný z toho dôvodu, že triangulácia funguje len na 2D rovine, a súradnicu Z sa pokúšali len triviálne odhadnúť [3, s. 9]. Výpočet bol teda primárne zameraný len na súradnice X a Y , ako je aj cieľom našej práce. Autori mali oproti našim predpokladom výhodu, že použitie GSW údajov obsahuje presné a rozsiahle informácie, od natočenia kamery vzhľadom na vertikálnu a horizontálnu os až po samotné súradnice (a mnoho iných, pre naše účely irelevantných, avšak v ich prípade využitých údajov). Natočenie kamery v čase zaznamenania snímky budeme musieť aproximovať len z polohy, ak takúto informáciu nezískame priamo hardvérovo.

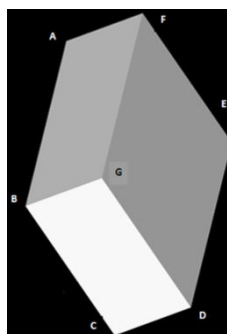
Možnosť lokalizovať objekt len pomocou jednej kamery bola demonštrovaná v článku [4], kde autori použili predpoklad, že polohy štyroch bodov terénu sú vopred známe. Na základe tohto sú schopní odvodiť informáciu o polohách ostatných bodov v priestore. Avšak naša metóda takýto predpoklad nepripúšťa.



Obrázok 1-12 Ilustrácia štyroch rôznych pohľadov pri procese lokalizácie objektu. [4, s. 5]

Lokalizovanie objektov je možné použiť spoločne s trasovaním objektu, ktoré je tiež jedným z doplnkových cieľov tejto práce v rámci výskumu, avšak v tomto prípade je použitie presné opačné tomu, aké zamýšľame my. Autori článku [5] sa zaoberali trasovaním vybraného objektu ako aj odhadovaním jeho polohy. Ich úlohou bolo trasovať a neskôr odhadovať polohu, my naopak primárne odhadujeme polohu a myšlienka je vylepšiť odhad trasovania na základe tejto polohy. Ich metóda má vyššie nároky na hardvér. Ich prístup uvádzame pre zmienku aj o možnosti odhadovať pozíciu pri trasovaní objektu.

Vhodnú ideu poskytli v publikácii [6], v ktorej preukázali, že získanie 3D polohy objektu z 2D obrázku je možné v prípade, že jeho rozmery a tvar sú známe. Pri takomto množstve informácií je možné počet tvarov, ktoré by boli projektované na 2D obrázok v takejto forme, podstatne zredukovať.

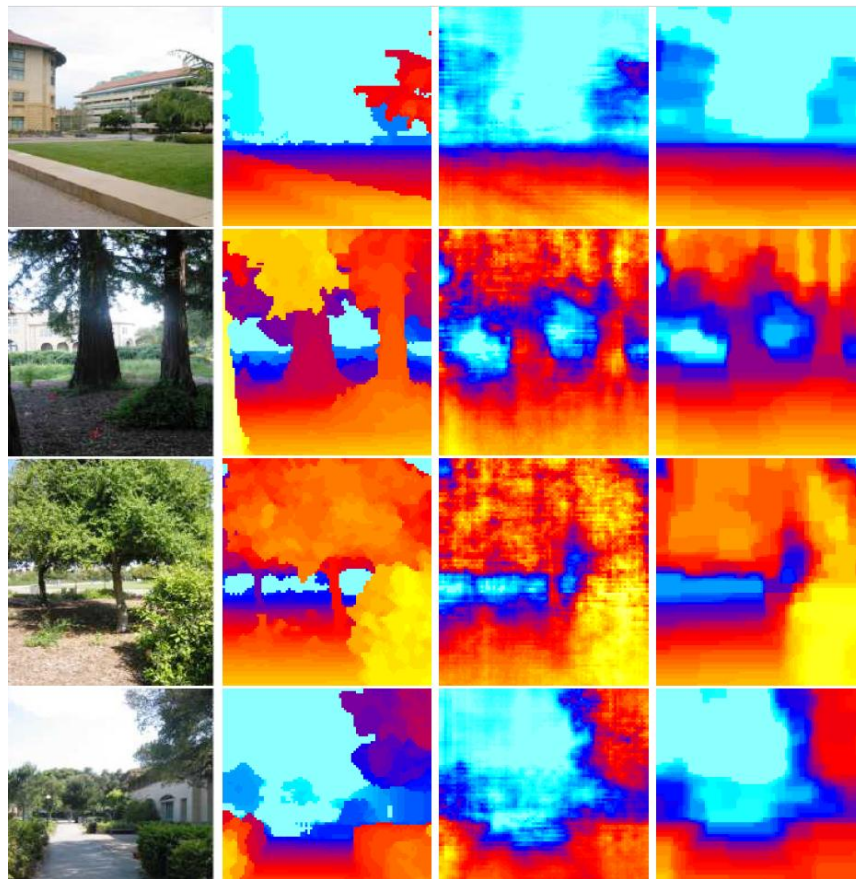


Obrázok 1-13 Pri znalosti niekoľko strán objektu je pozícia vypočítateľná.[6, s. 3]

Tento námet by bolo možné využiť pri našej metóde, kedy by sme predpokladali aj rozmery dopravného značenia, na ktoré je kladený štandard, a možnosti sú teda značne obmedzené.

Je evidentné, že fundamentálnym základom jednotlivých metód je triangulácia v nejakej forme (aj keď je často ukrytá len na úrovni elementárnych rovníc). Takýto prístup je aj laicky pochopiteľný, pretože sa jedná o pokus hľadať informácie o hĺbke o spoločnom objekte pre dve snímky, a pomyslený trojuholník sa v takejto situácii vytvorí vždy.

Snahu obísť trianguláciu preukázali autori v článku [7], ktorí použili metódy strojového učenia (angl. *machine learning*), aby naučili počítač pre daný vstup odhadovať HM. Technikou tzv. učenia s učiteľom (angl. *supervised learning*) trénovali model na vopred vytvorenom *datasete*, pre snímky ktorého boli hĺbkové mapy presne určené.

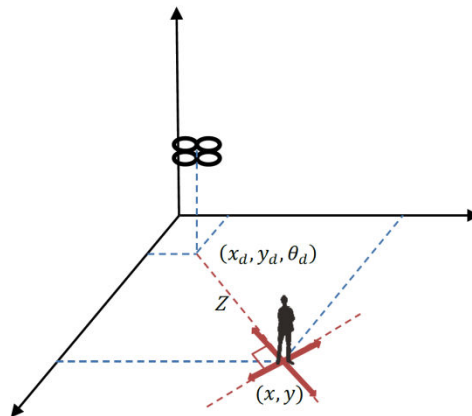


Obrázok 1-14 Ukážka hĺbkovej mapy pre vybrané obrázky [7, s. 7].

Vnímanie hĺbky nie je lokálna záležitosť, preto museli aplikovať postup, ktorý zahŕňal lokálne aj globálne vlastnosti (angl. *features*). Ukázali, že aj v neštruktúrovaných scénach je možná obnova hĺbkových máp v použiteľnej miere [7, s. 1].

Ako bolo spomínané, z matematického hľadiska je irelevantné, či je kamera stacionárna a objekt mobilný, alebo opačne. V dnešnej dobe narastajúceho počtu malých, bezpilotných lietadiel (angl. *drone*), ktoré sú často vybavené kamerou a rôznymi doplnkovými senzormi, odhad pozície objektu je potrebný aj v tejto oblasti.

Takýto *drone* je schopný lokalizovať chodcov, autá, prípadne môže byť použitý na samotné trasovanie nejakého vozidla [8, s. 1].



Obrázok 1-15Lokalizácia chodca z troch pozícií kamery. [8, s. 2]

Metóda, ktorá bola popísaná v [8] je založená na predvídateľnom vzore pohybu bezpilotného lietadla, ktorý je z matematického hľadiska pre lokalizáciu vhodný. Neskôr sú namiesto bežne používaných dvoch snímok využité až tri, na základe ktorých je určená pozícia objektu záujmu [8, s. 1].

2 Metodika práce, metody skúmania a návrh riešenia

Proces výskumu vhodnej metódy bol rozdelený do niekoľkých fáz. Prvotnou myšlienkou, ktorú bola snaha realizovať, bol výpočet vzdialenosti objektu od kamery. Táto informácia je nevyhnutná pre samotnú lokalizáciu, preto bolo za cieľ začať od základov.

Situácia nás donútila fázu výpočtu vzdialenosti rozdeliť na pohyb priamočiary voči objektu, a pohyb všeobecný. Je samozrejmosťou, že tieto situácie predstavujú rôznu úroveň zložitosti.

Po vytvorení matematického aparátu umožňujúceho odhad vzdialenosti nastal krok konštruovania matematických závislostí pre odhad samotnej pozície objektu.

Avšak ešte pred tým ako popíšeme jadro našej práce tak popíšeme dáta, na ktorých sme naše hypotézy testovali.

2.1 Vytváranie syntetického *datasetu*

Pri vývoji metód počítačového videnia je často problematické odhadnúť, či nepresnosti vo výsledkoch sú zapríčinené zlým postupom alebo inherentnými chybami v dátach, na ktoré sme postup aplikovali.

Techniky fotogrametrie vyžadujú pri overovaní správnosti presnosť kamery a údajov dokonale určenú. Takáto situácia je ideálna a v reálnom svete nedosiahnuteľná. Vzhľadom na tento fakt bolo nutné vytvoriť *dataset* syntetický, ktorý predstavuje ideálnu situáciu v umelých (laboratórnych) podmienkach, na ktorých môžeme naše metódy testovať a kalibrovať predtým, ako budú nasadené na *dataset* reálny.

Okrem vytvorenia ideálnych dát vzhľadom na to, že neobsahujú žiadne chyby, je rovnako možné vytvoriť aj ideálny *dataset*, ktorý obsahuje konkrétnu chybu, a my môžeme opäť objektívne posudzovať dosiahnuté výsledky. Takéto dôvody nás motivovali k vytvoreniu SD pomocou nástroja, ktorý sa používa pre vytváranie 3D scén, nazývaný Blender [16].

Potrebnými údajmi sú videosekvencia snímok s anotovaným objektom záujmu, parametre kamery ako je rozlíšenie, uhlové horizontálne uhlové zorné pole a pod. Poloha kamery spoločne s jej smerovaním sú poskytnuté rovnako. Samozrejým predpokladom je

snaha vytvoriť scénu, ktorá svojimi vlastnosťami bude reflektovať svet reálny v najväčšej možnej miere.

Jednotlivé metódy, ktoré budú na takomto *datasete* testované nie vždy vyžadujú všetky dostupné informácie. Ako bolo už uvedené, minimalistické predpoklady na funkčnosť metódy sú jedným z cieľov tejto práce.

Tak ako bolo neraz spomenuté, úloha je konkretizovaná na príklad pohybujúceho sa vozidla. Kamerou umiestnenou na vozidle je snaha zaznamenať objekt, ktorého polohu je cieľom lokalizovať. Týmto objektom je dopravná značka umiestnená pri ceste na pravej strane, ako je možné nájsť vo väčšine krajín sveta.

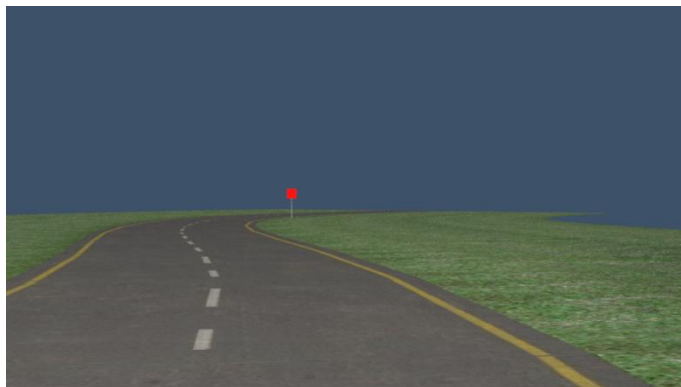
Teraz uvedieme podrobný popis SD, ktorými disponujeme. V kapitole dedikovanej testovaniu sa na jednotlivé *datasety* budeme odvolávať pomocou zavedených názvov, ktoré sú spoločne s parametrami *datasetu* uvedené v tabuľke nižšie.

Stĺpce s názvom *FoV (h/v)* označuje horizontálne a vertikálne uhlové zorné pole kamery respektíve. Počet snímok predstavuje počet tých snímok, na ktorých je objekt viditeľný, čiže pri prehliadke *datasetu* je možné nájsť preskočené miesta, nakoľko boli z *datasetu* vyradené z dôvodu nepotrebnosti. Všetky *datasety* tejto kategórie obsahujú len jeden objekt.

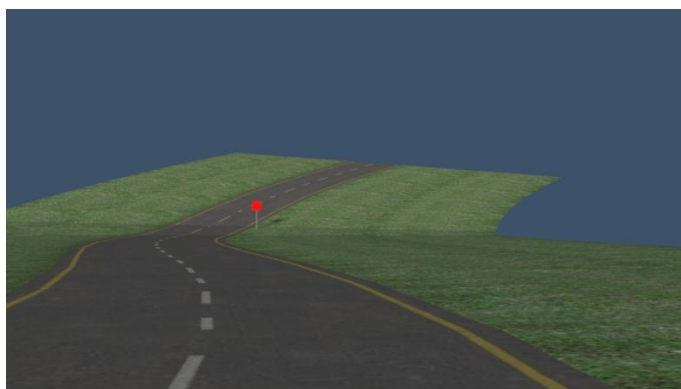
Názov	Kopce	Konštantné prevýšenie	Zakrivená cesta	FoV (h/v)	Rozlíšenie	Počet snímok
flat	nie	nie	áno	49.134 / 28.842	1280×720	254
hill	áno	nie	áno	49.134 / 28.842	1280×720	130
up	nie	áno	áno	49.134 / 28.842	1280×720	255
straight-left	nie	nie	nie	49.134 / 28.842	1280×720	32
straight-right	nie	nie	nie	49.134 / 28.842	1280×720	33

Tabuľka 1 Popis parametrov SD.

Ilustrácia vybraného úseku pre uvedené *datasety* v poradí ich vymenovania nasleduje ďalej.



Obrázok 2-1 Zakrivená cesta, bez prevýšenia, rovný terén (*dataset flat*).



Obrázok 2-2 Zakrivená cesta, kopcovitý terén (*dataset hill*).



Obrázok 2-3 Zakrivená cesta, konštantné stúpanie (*dataset tup*).

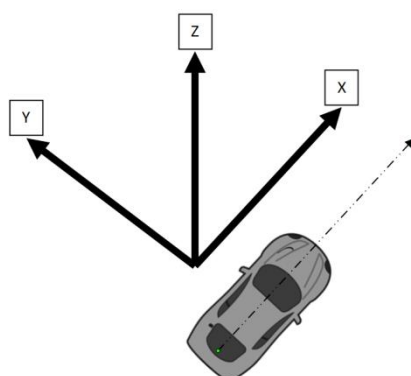


Obrázok 2-4 Rovná cesta, dopravné značenie umiestnené na ľavej strane (*dataset straight-left*).



Obrázok 2-5 Rovná cesta, dopravné značenie umiestnené na pravej strane (*dataset straight-right*).

Súradnicový systém nášho SD je nasledovný:



Obrázok 2-6 Súradnicový systém SD.

Objekt záujmu, v tomto prípade dopravné značenie, ktorého poloha je vždy blízko počiatku súradnicového systému. Auto sa pohybuje v smere rastúcej súradnice x , pri pohybe vľavo je súradnica y rastúca, a pohyb dohora predstavuje nárast pre súradnicu z .

SD bol vytvorený ako množina snímok s ich korešpondujúcimi dátovými súbormi, ktoré obsahujú všetky potrebné informácie. Informácie sú dostupné v JSON formáte. Nižšie uvádzame príklad takéhoto súboru. Je nutné poznamenať, že všetky *datasety* sú kompatibilné (spoločne s *datasetom* reálnym uvedeným neskôr) vzhľadom na formát údajov. Pre príklad uvedieme snímku číslo 247 z *datasetu flat*.



Obrázok 2-7 Snímka 247 z *datasetu flat*.

Korešpondujúci dátový súbor má nasledovný obsah v JSON formáte.

```
{
  "camera": {
    "azimuth": -17.650000000000006,
    "field_of_view": {
      "horizontal": 49.134,
      "vertical": 28.842
    },
    "position": {
      "x": -16.048,
      "y": -4.273,
      "z": 1.5
    }
  },
  "image": {
    "height": 720,
    "width": 1280
  },
  "objects": [
    {
      "bounding_box": {
        "height": 71,
        "width": 75,
        "x": 1119,
        "y": 290
      },
      "id": 1,
      "position": {
        "x": 0.0,
        "y": -5.0,
        "z": 1.87
      }
    }
  ]
}
```

Azimut kamery, horizontálne a vertikálne zorné pole a pozícia kamery v priestore.

Rozmery obrázku (snímku).

Zoznam viditeľných objektov.

X a Y súradnica spoločne so šírkou a výškou ohraničujúceho obdĺžnika pre aktuálny objekt v zozname.

Unikátne identifikačné číslo objektu platné pre celý *dataset flat*.

X, Y a Z súradnice pre aktuálny objekt v zozname.

2.2 Vytváranie reálneho *datasetu*

RD bol vytvorený jazdou automobilom po meste Žilina, na ktorom bola upevnená kamera. Počas nahrávania videa bol zaznamenávaný aj GPS signál, ktorý bol prítomný v intervaloch približne jednej sekundy. Dostupné sú časové pečiatky s granularitou približne jednej sekundy (počet meraní za sekundu je minimálne 1 a maximálne 2) spoločne so zemepisnou šírkou (angl. *latitude*) a zemepisnou dĺžkou (angl. *longitude*) vo formáte desatinných stupňov (angl. *decimal degrees*) *DDD.DDDDDD°*.

Zemepisná šírka a dĺžka sú poskytnuté s presnosťou na 6 desatinných miest. Vzhľadom aj na časovú pečiatku s presnosťou na jednu sekundu už dopredu avizujeme potenciálne signifikantný vplyv na presnosť výsledkov (ktorý sa budeme snažiť určiť pri testovaní). Takáto situácia nás núti údaje interpolovať, nakoľko počet snímok za sekundu pri video zázname je podstatne vyšší ako počet čerstvo nameraných informácií o polohe kamery.

Pre 6 desatinných miest súradníc je možné teoreticky vypočítať maximálnu možnú presnosť (vzhľadom na nejaké predpoklady). Historicky, meter bol z počiatku definovaný Francúzmi v roku 1793 ako $1/10^7$ zo vzdialenosti od severného pólu po rovník [19, s. 42]. Z tohto vyplýva, že 1° zemepisnej šírky pokrýva približne $10^7/90 = 111\,111, \bar{1}$ metrov. Ak urobíme predpoklad, že Zem je presná guľa a nie sploštený sféroid, potom rovnaká úvaha o presnosti platí aj pre zemepisnú dĺžku. Faktom však zostáva, že sploštenie Zeme predstavuje zanedbateľný vplyv. Nakoľko naše GPS údaje boli zachytené s presnosťou na 6 desatinných miest, potom teoretická presnosť na jeden GPS záznam je daná ako $111\,111/10^6 = 0,111111$, čo je približne $1/9$ metra.

Zariadenie, ktoré zaznamenávalo GPS poskytlo údaje uvedené nižšie. Uvádzame len niekoľko vybraných záznamov pre ukážku.

Čas	Zemepisná šírka	Zemepisná dĺžka	Prevýšenie	Presnosť	Smerovanie (vzhľadom na sever)	Rýchlosť
06:51:51	49.213488	18.739263	393	3	0	0
06:53:08	49.21329	18.737543	398	3	264.100006	9.33
06:53:33	49.21385	18.73742	401	3	63.799999	5.77
06:54:15	49.214101	18.746074	396	3	89.699997	15.1
06:55:07	49.213716	18.748354	403	3	114.5	1.66

Tabuľka 2 Ukážka GPS záznamov.

Videozáznam bol vytvorený pomocou mobilného zariadenia umiestneného na palubnej doske. *Dataset* teda obsahuje primárne GPS záznamy spoločne s video súborom.

Pri vytváraní *datasetu* nastali komplikácie po nejakom čase, pretože inicializácia GPS zariadenia trvala relatívne dlho, a na základe tohto nie je GPS záznam korektne synchronizovaný s video záznamom. Túto diskrepanciu bolo potrebné vyriešiť jednoduchým posunom času, ktorý musel byť odhadnutý z pozorovaní, a rovnako pripúšťame, že nesie so sebou malé riziko nepresnosti.

Tento fakt ale uvádzame z dôvodu, že presnosť je prvoradej dôležitosti v našej práci, a vzhľadom na nepriaznivé okolnosti je istá dávka nepresnosti už v reálnych dátach, ktorými disponujeme. Replikácia nahrávania *datasetu* už nebola vzhľadom na isté organizačné obmedzenia opätovne možná.



Obrázok 2-8 Ukážka RD – oblasť Žilina, blízko UNIZA.



Obrázok 2-9 Ukážka RD – oblasť Žilina, blízko UNIZA.

Takto vytvorené dáta boli neskôr anotované poloautomatickým spôsobom, nakoľko reálne situácie bežne obsahujú väčší počet dopravných značení, ktoré sú viditeľné na danej snímke. Anotačný nástroj vyprodukoval ďalší súbor (okrem kompletného priečinku jednotlivých snímok videa), ktorý obsahuje ohraničujúce obdĺžniky pre niektoré vybrané

snímky spoločne s oficiálnou kategóriou dopravného značenia. Nižšie uvádzame úvodnú časť pôvodného XML súboru, ktorý je výstupom z anotačného nástroja.

```
<?xml version="1.0" encoding="utf-8"?>
<data>
<metadata>
<data_id>sa_dataset</data_id>
<parent></parent>
<version_major>2</version_major>
<xml_sid></xml_sid>
<description>Anotacny nástroj v1.02</description>
</metadata>
<images>
<image>
<src>20150617_085226_FHD/20150617_085226_00631.jpg</src>
<boundingboxes></boundingboxes>
</image>
<image>
<src>20150617_085226_FHD/20150617_085226_00632.jpg</src>
<boundingboxes>
<boundingbox>
<x_left_top>76</x_left_top>
<y_left_top>327</y_left_top>
<width>58</width>
<height>58</height>
<class_name>
<class_id>B1</class_id>
</class_name>
</boundingbox>
</boundingboxes>
</image>
```

Dataset v takomto formáte bolo potrebné previesť do nášho interného formátu pre tento projekt, aby bol zosúladený aj s formátom predstaveným v predchádzajúcej podkapitole ohľadom SD. Avšak ešte pred samotným prevodom bolo nutné získať skutočnú zemepisnú šírku a dĺžku dopravných značení, ktoré sa v *datasete* nachádzajú. Opäť je nutné poznamenať, že nie všetky dopravné značenia sme použili. Jedným

z dôvodov bol aj spomenutý problém so synchronizáciou, čiže istú časť video záznamu bolo potrebné vynechať. Ďalším dôvodom bola nekorektná detekcia dopravného značenia, nakoľko bola vykonávaná poloautomaticky.

Problémy nastávali pri zisťovaní skutočnej polohy dopravného značenia. Túto úlohu sme sa pokúsili v začiatkoch vyriešiť manuálnym meraním GPS polohy použitím mobilného telefónu. Úloha bola v princípe jednoduchá: priložiť mobilný telefón k podpernej tyči dopravného značenia, použiť niekoľko nezávislých mobilných aplikácií umožňujúcich získať aktuálnu GPS polohu, zaznamenanie tejto polohy a neskôr tieto dáta manuálne spracovať do požadovaného formátu.

Na naše prekvapenie bolo toto riešenie použiteľné len v niektorých prípadoch, pretože GPS v mobilnom telefóne bolo pri viacej pokusoch veľmi nepresné, a polohu sme získali len orientačne, ale pre naše účely je nutné polohu poznať čo najpresnejšie.

Riešením bolo použitie *Google Maps* (GM) v kombinácii s *Google Street View* (GSV). „Od oka“ sme odhadli možnú pozíciu dopravného značenia v priestore s maximálnym priblížením, ktoré GSV umožňoval. Následne sme sa premiestnili do GM pohľadu, a vyhľadali zemepisnú šírku a výšku pixlu, na ktorý sme klikli. Takto nájdenú hodnotu sme prehlásili ako možnú pozíciu objektu.



Obrázok 2-10 Ukážka procesu odhadovania súradníc konkrétneho miesta pomocou GM.

Vzhľadom na to, že značenie je často prítomné pri nejakom zachytanom bode akým môže byť napríklad priechod pre chodcov, ktorý je z GM pohľadu dobre viditeľný, vo väčšine prípadov máme istú dávku dôvery v takto namerané pozície, nakoľko aj pri dodatočnej validácii inými nástrojmi schopnými renderovať body na mape na základe dodaných pozícií sme nenašli relevantné odchýlky. Avšak našli sa dopravné značenia, pri ktorých je naša dôvera v pozíciu redukovaná, nakoľko sa nachádzali na miestach ako sú rovné cesty, kde nie sú prítomné ani stromy, priechody pre chodcov alebo križovatky a podobné pomocné prvky, zreteľne viditeľné z pohľadu GM. Na základe tohto sme sa rozhodli, že pri testovaní budú objekty RD rozdelené do dvoch skupín odzrkadľujúc našu dôveru v ich skutočnú pozíciu, a budú analyzované separátne. GM rovnako poskytuje súradnice s presnosťou na 6 desatinných miest.

Poloautomatická anotácia spôsobila, že BB jednotlivých dopravných značení nebol prítomný na všetkých snímkach, na ktorých sa daný objekt vyskytoval. Na základe tohto bolo potrebné dáta interpolovať, a aplikovať veľmi triviálne trasovanie objektov.

Prítomnosť oficiálnej kategórie dopravného značenia bola značnou výhodou, pretože v kombinácii s riedkou prítomnosťou BB bolo možné nájsť, kde by sa teoreticky mohol BB nachádzať na snímkach, ktoré BB neobsahujú a mali by.

Nech t_i predstavuje bod v čase, kedy je známy BB objektu a nech t_j je jeden z nasledovných bodov v čase (ideálne najbližší možný), kedy sú známe rovnaké informácie. Je zrejmé, že $t_i < t_j$. Uvažujme situáciu, kedy existuje čas t_k , taký že $t_i < t_k < t_j$ a cieľom je odhadnúť, aká je pozícia BB v čase t_k , keď je táto pozícia známa v časoch t_i a t_j .

Ak predpokladáme triviálny, lineárny model, kedy transformácia tvaru BB z času t_i do t_j prebieha priamočiara, potom je možné aplikovať základnú lineárnu interpoláciu medzi dvomi vektormi v priestore. Nech $\mathbf{b}_i = (x_i y_i w_i h_i)$ je BB v čase t_i , a analogicky $\mathbf{b}_j = (x_j y_j w_j h_j)$ je BB v čase t_j . Úlohou je nájsť BB v čase t_k , ktorý je určený ako $\mathbf{b}_k = (x_k y_k w_k h_k)$. Transformáciu definujeme nasledovne [20, s. 2]:

$$\mathbf{b}_k = \left(\frac{t_j - t_k}{t_j - t_i} \right) \mathbf{b}_i + \left(\frac{t_k - t_i}{t_j - t_i} \right) \mathbf{b}_j \quad (2.1)$$

Je evidentné, že v dôsledku pohybu automobilu, napríklad zatáčania, zmeny rýchlosti, prítomnosti perspektívnej deformácie a mnoho iných faktorov takýto prechod nie

je v skutočnosti lineárny, avšak v našom prípade, kedy je frekvencia snímok 25 za sekundu a anotácie BB sú dostupné niekoľko krát za sekundu je takýto model vhodnou aproximáciou. Toto tvrdenie sme overili vizualizáciou BB a trasovaním, a na pohľad výsledky korešpondovali realite s minimálnymi odchýlkami. Zdôrazňujeme, že aj takáto aproximácia malou mierou prispieva k celkovej nepresnosti.

Po interpolácií BB a určení jednoznačných identifikačných čísel pre jednotlivé dopravné značenia, ktoré sa v *datasete* nachádzajú, prišiel na rad odhad GPS pozície pre konkrétne snímky. Vzhľadom na to, že frekvencia snímok ďaleko prevyšuje frekvenciu GPS meraní, opäť bolo potrebné určiť časovú pečiatku pre jednotlivé snímky a následne čo najpresnejšie odhadnúť, aká mohla byť poloha vozidla v danom čase.

Ak použijeme rovnaké značenie ako v úvahe vyššie, kedy t_i a t_j , také, že $t_i < t_j$, predstavujú časy, v ktorých sú údaje o zemepisnej šírke a výške dostupné ako $\mathbf{g}_i = (x_i y_i)$ a $\mathbf{g}_j = (x_j y_j)$ respektíve, potom chýbajúce údaje $\mathbf{g}_k = (x_k y_k)$ v čase t_k , $t_i < t_k < t_j$, určíme lineárnou interpoláciou obdobným spôsobom [20, s. 2]:

$$\mathbf{g}_k = \left(\frac{t_j - t_k}{t_j - t_i} \right) \mathbf{g}_i + \left(\frac{t_k - t_i}{t_j - t_i} \right) \mathbf{g}_j \quad (2.2)$$

V takomto prípade je lineárna aproximácia lepším modelom, pretože pohyb auta aj v zákrute je na časovom intervale pol sekundy veľmi podobný čiare, ale obraz v zákrutách vykazuje dramatickejšie zmeny a nepresnosti pri lineárnej interpolácii.

Jediný požadovaný parameter kamery pre naše prístupy je horizontálne uhlové zorné pole kamery (FoV horizontálne). Túto hodnotu je možné určiť aj empiricky na základe video záznamu, s nejakou úrovňou chybovosti, avšak je možné obstať špecifikáciu zariadenia, ktorým bol *dataset* vytvorený. Takáto špecifikácia by mala obsahovať ohniskovú vzdialenosť kamery pri jednotlivých úrovniach priblíženia ako aj rozmery senzora. Po prevode týchto dvoch parametrov na identické jednotky je možné použiť vzťah nižšie. Jeho odvodenie nebudeme uvádzať nakoľko nie je relevantné, len spomenieme, že sa jedná o fundamentálny vzťah z oblasti optiky.

Nech w a h sú šírka a výška senzora a f je ohnisková vzdialenosť kamery v rovnakých jednotkách dĺžky ako w a h . Potom pre výpočet horizontálneho zorného pola (V_h) platí:

$$V_h = 2 \tan^{-1} \left(\frac{w}{2f} \right) \quad (2.3)$$

A pre výpočet vertikálneho zorného pola (V_v) platí:

$$V_v = 2 \tan^{-1} \left(\frac{h}{2f} \right) \quad (2.4)$$

V tejto fázy disponujeme všetkými potrebnými údajmi aby sme mohli ukončiť spracovanie *datasetu* a vygenerovať RD vo formáte, ktorý je kompatibilný s našou testovacou aplikáciou a SD.

Teraz definujeme parametre, ktorých význam je identický ako v prípade SD.

Názov	Kopce	Konštantné prevýšenie	Zakrivená cesta	FoV (h/v)	Rozlíšenie	Počet snímok	Počet objektov
zilina	mierne	nie	áno	59.52084 / 46.60603	1920 × 1080	1474	37

Tabuľka 3 Parametre *datasetu zilina*.

Výsledný obsah JSON súboru je pre ukážku poskytnutý ďalej. Jedinou zmenou voči popisu SD je rozdielna notácia polohy kamery a dopravného značenia, nakoľko v tomto prípade je použitá zemepisná šírka a zemepisná dĺžka.



Obrázok 2-11 Snímka 112 z *datasetu zilina*.

```
{
  "camera": {
    "azimuth": 113.61201011591541,
    "field_of_view": {
```

```
        "horizontal": 60.0,
        "vertical": 45.0
    },
    "position": {
        "latitude": 49.211751320004446,
        "longitude": 18.754854874634347
    }
},
"image": {
    "height": 1080,
    "width": 1920
},
"objects": [
    {
        "bounding_box": {
            "height": 159,
            "width": 158,
            "x": 1728,
            "y": 73
        },
        "id": 1,
        "position": {
            "latitude": 49.211694,
            "longitude": 18.75476
        }
    }
],
"timestamp": 247778.6663701702
}
```

2.3 Vzdialenosť od objektu pri priamočiarom pohybe

Jeden z prvých, aj keď nie pre našu situáciu globálne aplikovateľných prístupov, ktorý sa nám podarilo vymyslieť, bol odhad vzdialenosti objektu od kamery len za použitia dvoch snímok a pomeru dĺžok strany daného objektu. Tento prístup dosahuje vynikajúce

výsledky pri priamočiarom pohybe, z čoho vyplýva aj nami vytvorený názov pre tento prístup.

Matematické závislosti, ktorých odvodenie je dôkladne popísané, sa zakladajú len na znalostiach matematiky v oblasti trigonometrie. Inšpirácia nepochádzala zo žiadneho konkrétneho vedeckého článku. Pripúšťame možnosť, že identické vzťahy boli vyvedené aj inými autormi, vzhľadom na to, že takáto situácia je pravdepodobná, nakoľko sa nejedná o komplexnú matematickú formuláciu.

Prístup uvádzame z dôvodu, že je aplikovateľný v prípade, kedy je objekt stacionárny, a kamera sa pohybuje len priamočiarym smerom k objektu. Takáto situácia môže nastať v oblasti robotiky, kedy jednoduché roboty poskytujú pohyb vždy len po jednej osi (pre pohyb iným smerom je nutné robota zastaviť, urobiť rotáciu, a následne iniciovať pohyb novým smerom). Rešpektujúc tieto reštrikcie, potom nám postačuje krátky pohyb v smere voči objektu, a na základe získaných informácií je možné vypočítať vzdialenosť k objektu záujmu.

Tento postup zlyháva, ak sa kamera pohybuje voči objektu všeobecne, čiže sa menia súradnice dvoch alebo až troch osí súčasne po dlhšiu dobu.

Pre účely demonštrácie použijeme výšku objektu ako atribút, ktorý budeme sledovať. V konečnom dôsledku na tomto nezáleží, prvoradá je, aby daný atribút bol jednorozmerný, čiže je možné použiť aj šírku, prípadne priemer kruhu a pod. Potreba jednorozmernej hodnoty je výhodou v tom, že prístup je jednoduchý, avšak nevýhodou vtedy, ak je pohyb po viacej osiach naraz, pretože dochádza k extrémnej chybovosti. Myšlienku sa nám nepodarilo rozšíriť do viacej rozmerov, ani formálne ukázať, že niečo také nie je možné.

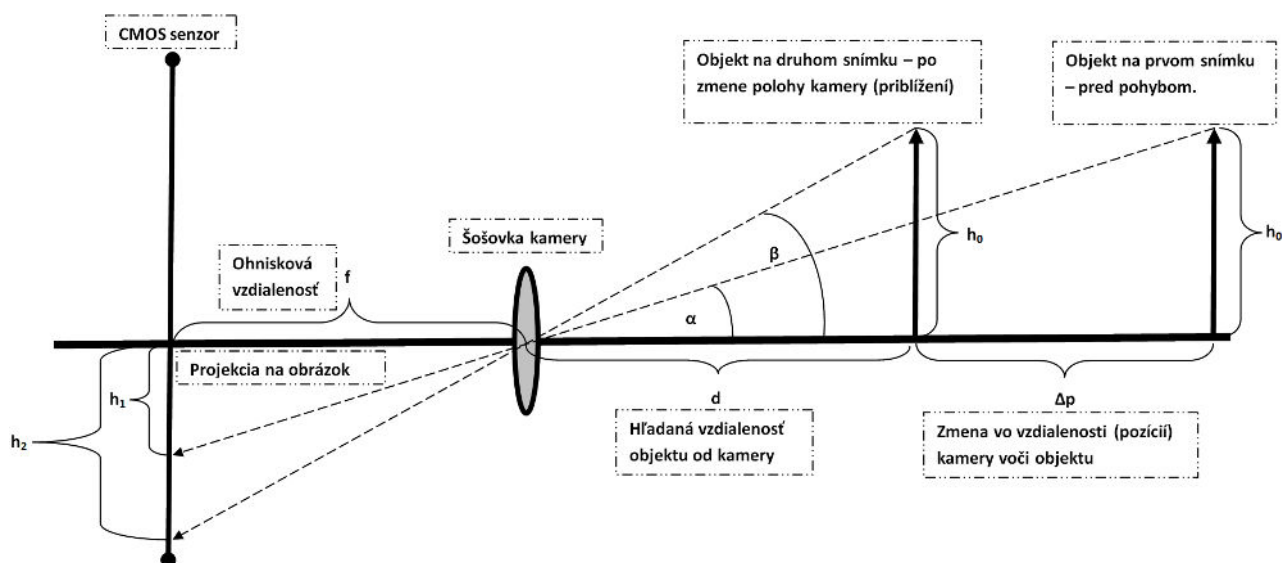
Metóda pre svoje fungovanie potrebuje len dve snímky a informáciu o zmene polohy, čiže vzdialenosť, o ktorú sa buď kamera priblížila k objektu, alebo objekt ku kamere. Vnútorne parametre kamery nie sú vôbec potrebné, ako aj skutočná veľkosť objektu.

Nech h_1 je výška objektu projektovaná na obrázok meraná v pixloch na prvom obrázku, a h_2 je výška objektu v pixloch na druhom obrázku. Predpokladáme, že pohyb kamery je smerom k objektu, čiže $h_1 < h_2$. Zdôrazňujeme, že ostrá nerovnosť je nutnosťou, nakoľko rozdiel musí byť badateľný. Nech Δp určuje zmenu vo vzdialenosti

v ľubovoľných jednotkách. Pre vzdialenosť objektu od kamery, d , vzhľadom na druhý snímok, ktorá bude v určená v jednotkách rovnakých ako Δp , platí nasledovný vzťah:

$$d = \frac{\Delta p}{\frac{h_2}{h_1} - 1} \quad (2.5)$$

Pri známej zmene vo vzdialenosti je potrebný len pomer niektorej z vybraných strán objektu na obrázku v pixloch. Tento vzťah nám určí vzdialenosť objektu od kamery vzhľadom na druhý snímok, teda ten, keby bol objekt bližšie. Nižšie uvádzame detailné odvodenie daného vzťahu. Čitateľovi odporúčame nasledovať ilustračný obrázok.



Obrázok 2-12 Ilustrácia odvodenia vzťahu pre výpočet vzdialenosti.

Nech h_0 je skutočná výška pomysleného objektu, v ľubovoľných jednotkách, a označenia h_1 , h_2 , Δp a d majú ponechaný význam. Dodatočne zavádzame f ako ohniskovú vzdialenosť kamery, spoločne s uhlami α a β reprezentujúcimi uhol, pod ktorým lúč z vrcholu objektu dopadá na šošovku kamery. Model je inšpirovaný klasickým fungovaním tzv. *pinhole camera*.

Z obrázku vyplýva nasledovná sústava dvoch (resp. štyroch) rovníc, ktoré budú neskôr upravené, aby vyjadrili d ako závislú premennú od h_1 , h_2 a Δp .

$$\tan \alpha = \frac{h_0}{d + \Delta p} = \frac{h_1}{f} \quad (2.6)$$

$$\tan \beta = \frac{h_0}{d} = \frac{h_2}{f} \quad (2.7)$$

Pre f z rovnice 2.7 platí:

$$f = \frac{d \cdot h_2}{h_0} \quad (2.8)$$

Po dosadení vzťahu 2.8 za f v rovnici 2.6:

$$\frac{h_0}{d + \Delta p} = \frac{h_1 \cdot h_0}{d \cdot h_2} \quad (2.9)$$

$$d \cdot h_2 \cdot h_0 = (d + \Delta p) \cdot h_1 \cdot h_0 \quad (2.10)$$

$$d \cdot h_2 = (d + \Delta p) \cdot h_1 \quad (2.11)$$

$$d \cdot h_2 = d \cdot h_1 + \Delta p \cdot h_1 \quad (2.12)$$

$$d \cdot h_2 - d \cdot h_1 = \Delta p \cdot h_1 \quad (2.13)$$

$$d \cdot (h_2 - h_1) = \Delta p \cdot h_1 \quad (2.14)$$

$$d = \frac{\Delta p \cdot h_1}{h_2 - h_1} \quad (2.15)$$

$$d = \frac{\Delta p}{\frac{h_2}{h_1} - 1} \quad (2.16)$$

2.4 Spoločný základ pre lokalizáciu z dvoch snímok

Opäť je potrebné zdôrazniť, že nižšie uvedené matematické vzťahy boli vyvedené v procese riešenia nášho problému a následného experimentovania, preto nebudú vždy uvedené odkazy na relevantné vedecké publikácie v tejto oblasti, aj keď nezamietame možnosť existencie veľmi podobných alebo priam identických prístupov. S jedným sme sa stretli v článku [3], kde autori použili trianguláciu pomocou sínusovej vety na lokalizáciu objektu, čo sa podarilo odvodiť aj nám.

2.4.1 Predpoklady

Dve metódy uvedené v nasledujúcich dvoch kapitolách vyžadujú na vstupe dve snímky z videozáznamu. Samozrejým predpokladom je prítomnosť objektu záujmu na oboch snímkach. Ideálna situácia je, aby na oboch daných snímkach bol badateľný, ale nie moc veľký rozdiel v grafickej reprezentácii objektu. Inak povedané, ideálne je, aby na jednotlivých snímkach bol objekt umiestnený v tak rozdielnych vzdialenostiach od kamery aby bol zabezpečený dostatočný rozdiel vo veľkosti BB. Takto vágne uvedená požiadavka bez konkrétnych hodnôt môže pôsobiť máľúco a preto sa ju pokúsime ozrejmiť.

Rozlíšenie kamery má vplyv na presnosť výpočtu, a z tohto dôvodu ak pracujeme s obrazom s vysokým rozlíšením, už pri malej zmene v pozícií kamery bude možné na úrovni pixlov detegovať relevantné zmeny. Malé zmeny v polohe je možné detegovať presnejšie ak je objekt bližšie ku kamere. Pri nízkom rozlíšení alebo veľmi vzdialenom objekte je možné, že sa samotná kamera pohne o niekoľko metrov a stále objekt bude reprezentovaný veľmi podobným prípadne rovnakým BB. V takejto situácii metóda zlyháva. V oboch našich metódach ďalej spomenutých k zlyhaniu môžu prispievať aj iné dôvody, ktoré spomenieme priamo pri predstavení danej metódy.

Nakoľko obidve metódy potrebujú dva snímky, výpočet uhlu, pod ktorým je objekt viditeľný na danom snímku je opäť rovnaký. Teraz uvedieme spoločný základ, na ktorom budú rôznym spôsobom stavať jednotlivé metódy.

Nech f_1 a f_2 sú snímky video záznamu, pomocou ktorých je snaha určiť pozíciu objektu. Nech $t(f)$ je čas, kedy bol daný snímok f zhotovený, potom platí:

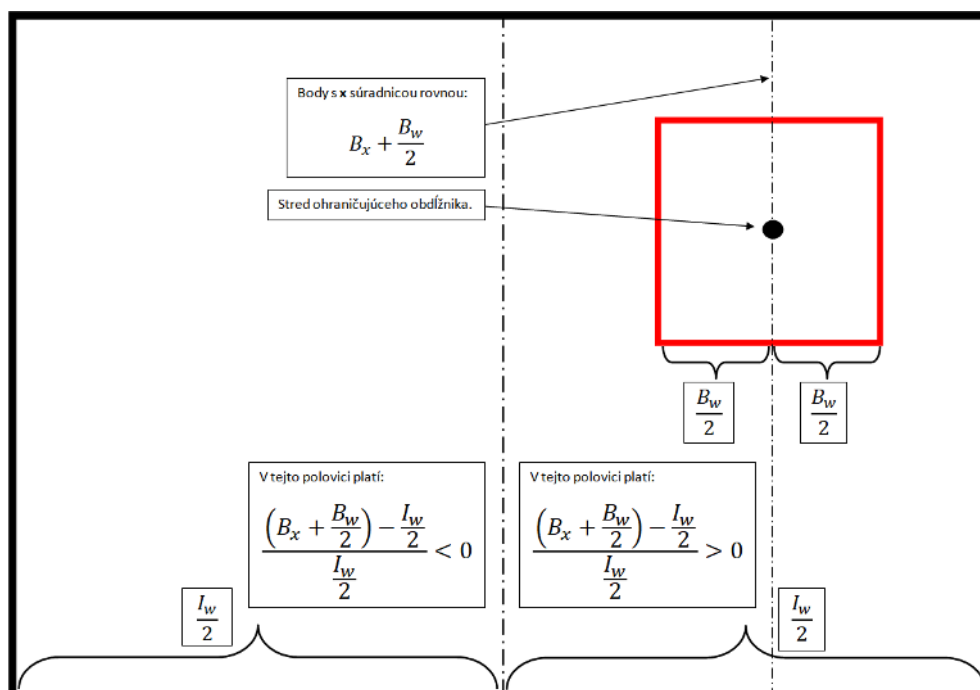
$$t(f_1) < t(f_2) \quad (2.17)$$

Azimut kamery pre daný snímok bude v našich výpočtoch reprezentovaný uhlom natočenia od pomysleného severu. *Dataset* túto hodnotu priamo obsahuje a teda jej výpočet nebude potrebný. Prístup by bolo možné aplikovať aj pri výpočte azimutu, tento výpočet by bol zahrnutý ako súčasť prípravy dát, ale pre jednoduchosť od tohto kroku abstrahujeme. Pre všeobecné použitie a hlavne testovacie účely je ideálne azimut vypočítat softvérovo a do *datasetu* jeho hodnotu zahrnúť a už viac ju nemodifikovať, ako sme učinili aj v našom prípade.

Nech α je uhol, pod ktorým je objekt viditeľný na snímke f_1 , a β uhol analogicky určený pre snímku f_2 . Tieto uhly sú vypočítané vzhľadom na smerovanie vozidla. Oba tieto uhly ležia na intervale $\langle -\frac{F}{2}, \frac{F}{2} \rangle$, kde F určuje zorné pole kamery. Uvedené matematické závislosti platia rovnako pri použití stupňov alebo radiánov.

Uhol, ktorého hodnota je záporná predstavuje situáciu, kedy je objekt viditeľný na ľavej strane, kladná hodnota reprezentuje pravú stranu, a nulová presný stred. Nech B_x a B_y reprezentujú x a y súradnicu ľavého horného rohu rozpoznaného dopravného značenia na danom snímku (popis BB objektu). B_w a B_h určujú šírku respektíve výšku tohto obdĺžnika a I_w a I_h platia identicky pre celý obrázok. Uhol α resp. β potom vypočítame nasledovne:

$$\alpha = \frac{F}{2} \left(\frac{\left(B_x + \frac{B_w}{2} \right) - \frac{I_w}{2}}{\frac{I_w}{2}} \right) \quad (2.18)$$



Obrázok 2-13 Ilustrácia výpočtu uhlu, pod ktorým je stred BB viditeľný.

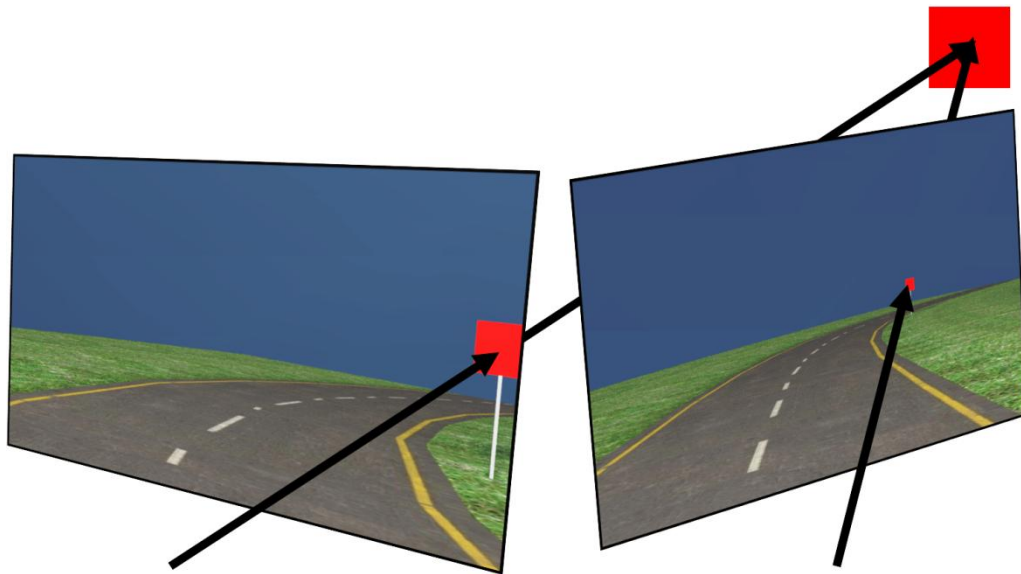
Pre naše účely výpočtu zavedieme pojem *azimut pixlu* (AP), ktorý je odvodený od azimutu samotného vozidla a dá sa interpretovať ako natočenie lúča od severu vyslaného z optického stredu kamery, ktorý prechádza daným pixlom, rovnako ako v prípade natočenia automobilu. Pri výpočte sa jedná len o pričítanie alebo odčítanie hodnoty α resp. β od aktuálneho azimutu vozidla. Pixel záujmu je v tomto prípade stred dopravného značenia.

Všetky zavedené značenia, pojmy a vzťahy budú využívané v popise pre nasledujúce dve metódy.

2.5 Určovanie polohy objektu na základe stretnutia priamok

V tejto časti uvedieme ďalšiu metódu, ktorú sme vytvorili za účelom lokalizácie objektu. Tento prístup vychádza z poznatkov o priamkach v priestore. Základná myšlienka spočíva v konštrukcii pomyslenej priamky (dá sa na ňu pozeráť ako na lúč), ktorá prechádza stredom kamery pre obidve snímky, na ktorých je objekt záujmu prítomný, a následne pretína daný objekt (v strede jeho BB). Na základe tohto je možné použiť len

dve snímky, skonštruovať spomínané priamky a vypočítať, kde sa dané priamky v priestore pretnú. Táto pozícia bude prehlásená ako náš odhad pozície daného objektu.



Obrázok 2-14 Vizualizácia lokalizácie pomocou dvoch pomyslených priamok.

Naša metóda je teda založená na výpočte bodu, v ktorom sa dve priamky pretnú. Tieto dve priamky sú skonštruované pomocou známej polohy kamery pre oba uvažované snímky a uhlu, pod ktorým je daný objekt viditeľný v rámci daného snímku (odpovedajú výpočtu uvedenému v predošlej kapitole).

Nech bod (x_1, y_1) je poloha kamery na snímku f_1 a bod (x_2, y_2) je poloha kamery na snímku f_2 . Nech p_1 a p_2 určujú AP na jednotlivých snímkach f_1 resp. f_2 . Cieľovú hodnotu x_t určíme nasledovne:

$$x_t = \frac{y_2 - y_1 + \tan(p_1) x_1 - \tan(p_2) x_2}{\tan(p_1) - \tan(p_2)} \quad (2.19)$$

Cieľovú hodnotu y_t vypočítame nasledovne:

$$y_t = \tan(p_1) x_t + y_1 - \tan(p_1) x_1 \quad (2.20)$$

Z výrazu v menovateli rovnice 2.19 je evidentné, že metóda má obmedzenia. Tento výraz je nedefinovaný pre prípad kedy platí: $\tan(p_1) - \tan(p_2) = 0$. Je nutné uviesť, že z výpočtového hľadiska je toto obmedzenie ešte striktniejšie, pretože pracujeme s reálnou aritmetikou počítača, a v takom prípade môže výpočet rezultovať v nedefinovaný výsledok keď $\tan(p_1) - \tan(p_2) \leq \varepsilon$, kde ε reprezentuje minimálny možný rozdiel dvoch reálnych čísel daného počítača, na ktorom je výpočet vykonávaný. Tento rozdiel by však nemal byť problémom pri implementácii a praktickej realizácii s cieľom získať použiteľné výsledky, nakoľko sa v prvom rade neodporúča používať dve snímky kedy by dochádzalo, že sú dané

priamky paralelné. Takáto situácia môže nastať keď je veľmi malý rozdiel v polohe kamery pri jednotlivých snímkach, a v takom prípade už nie je prítomný ani dostatočný rozdiel vo veľkosť BB, a o tomto už bola diskusia vyššie, prečo takýto prípad nie je vhodný. Takéto situácie skrátka odporúčame vo výpočte ignorovať.

Vzťahy vyššie uvedené boli odvodené nasledovne. Definujme dve základné rovnice, ktoré určujú, že dve nami nájdené priamky musia prechádzať cieľovým bodom (x_t, y_t) . Vychádzame zo štandardného zápisu všeobecnej rovnice $y = mx + b$.

$$y_t = m_1 x_t + b_1 \quad (2.21)$$

$$y_t = m_2 x_t + b_2 \quad (2.22)$$

Kde m_1 a m_2 predstavujú smernice nájdených dvoch priamok, a b_1 a b_2 body, v ktorých pretínajú os y . Je zrejmé, že smernica priamky je v tomto prípade určená hodnotou AP, preto je možné rovnice prepísať ako:

$$y_t = \tan(p_1) x_t + b_1 \quad (2.23)$$

$$y_t = \tan(p_2) x_t + b_2 \quad (2.24)$$

Z uvedeného vyplýva rovnosť:

$$\tan(p_1) x_t + b_1 = \tan(p_2) x_t + b_2 \quad (2.25)$$

Hľadané priamky sú definované ako:

$$y_1 = m_1 x_1 + b_1 \quad (2.26)$$

$$y_2 = m_2 x_2 + b_2 \quad (2.27)$$

Rovnosti nám poskytujú možnosť vyjadriť koeficienty b_1 a b_2 , ktoré budú potrebné pre výpočty neskôr.

$$b_1 = y_1 - m_1 x_1 = y_1 - \tan(p_1) x_1 \quad (2.28)$$

$$b_2 = y_2 - m_2 x_2 = y_2 - \tan(p_2) x_2 \quad (2.29)$$

Po dosadení do rovnice 2.25 (s použitím m namiesto $\tan x$ pre jednoduchosť):

$$m_1 x_t + (y_1 - m_1 x_1) = m_2 x_t + (y_2 - m_2 x_2) \quad (2.30)$$

$$m_1 x_t + y_1 - m_1 x_1 = m_2 x_t + y_2 - m_2 x_2 \quad (2.31)$$

$$m_1 x_t - m_2 x_t + y_1 - m_1 x_1 = y_2 - m_2 x_2 \quad (2.32)$$

$$m_1 x_t - m_2 x_t = y_2 - m_2 x_2 - y_1 + m_1 x_1 \quad (2.33)$$

$$m_1 x_t - m_2 x_t = y_2 - m_2 x_2 - y_1 + m_1 x_1 \quad (2.34)$$

$$x_t(m_1 - m_2) = y_2 - y_1 + m_1 x_1 - m_2 x_2 \quad (2.35)$$

$$x_t = \frac{y_2 - y_1 + m_1 x_1 - m_2 x_2}{m_1 - m_2} \quad (2.36)$$

$$x_t = \frac{y_2 - y_1 + \tan(p_1) x_1 - \tan(p_2) x_2}{\tan(p_1) - \tan(p_2)} \quad (2.37)$$

Pre výpočet y_t je možné vybrať jednu z dvoch rovníc 2.28 a 2.29 a dosadiť hodnoty x_t a b_1 resp. b_2 (v závislosti od výberu). My vyberieme prvú rovnicu. Potom pre y_t platí:

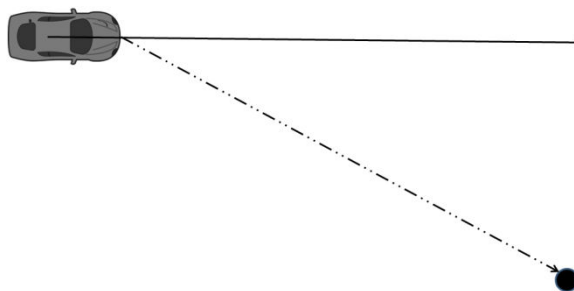
$$y_t = \tan(p_1) x_t + y_1 - \tan(p_1) x_1 \quad (2.38)$$

2.6 Určovanie polohy objektu na základe triangulácie

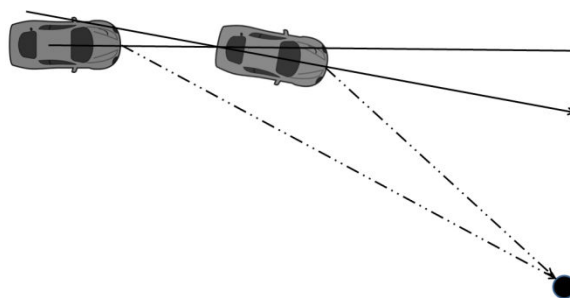
Pre krátke pripomenutie už uvedeného, metódu principiálne identickú našej je možné nájsť v článku [3], v ktorom rovnako použili trianguláciu na základe sínusovej vety pre určenie polohy objektu. Ich aplikácia bola však rozdielna od našej, nakoľko lokalizovali polohy pomocou GSW.

O triangulácií je možné hovoriť aj v prípade metódy predošlej, ale v tom prípade to nie je úplne presné, nakoľko nijakým spôsobom nevyužívame vlastnosti trojuholníka aby tvorili základ celého postupu (aj keď ho je v našej situácii možné bez problémov skonštruovať). Tento základ však tvoria priamky. V tomto prípade je ale situácia iná, a my skutočne vytvárame pomyslený trojuholník v skúmanej scéne a na základe jeho vlastností kalkuluje potrebné informácie.

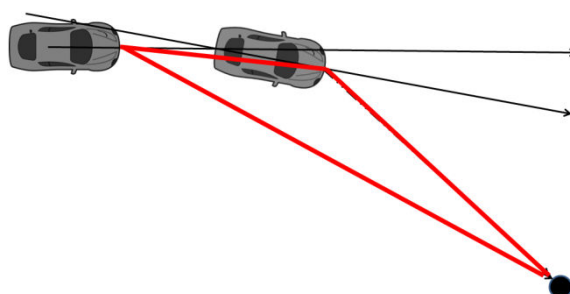
Nižšie uvádzame 3 obrázky demonštrujúce zostrojenie trojuholníka, pomocou ktorého vypočítame pozíciu objektu. Nech objekt reprezentovaný čiernou bodkou v pravom dolnom rohu je dopravné značenie. Čiara prechádzajúca vozidlom je smerová priamka určujúca jeho azimut, a prerušovaná čiara určuje uhol, pod ktorým je objekt viditeľný.



Obrázok 2-15 Situácia na prvej snímke.



Obrázok 2-16 Situácia na druhej snímke.



Obrázok 2-17 Konštrukcia pomysleného trojuholníka v danej scéne.

Nech bod (x_1, y_1) je poloha kamery na snímke f_1 a bod (x_2, y_2) je poloha kamery na snímke f_2 . Nech p_1 a p_2 určujú AP na jednotlivých snímkach f_1 resp. f_2 . V tomto prístupe je potrebná aj smernica priamky, ktorá spojuje dva body určené polohou kamery v jednotlivých časoch. Nech m_0 predstavuje smernicu priamky, ktorá prechádza bodmi (x_1, y_1) a (x_2, y_2) . Pre m_0 platí:

$$m_0 = \frac{y_2 - y_1}{x_2 - x_1} \quad (2.39)$$

Nakoľko budeme pracovať so smernicami troch priamok, pre jednoduchosť zavedieme smernicu zvyšných dvoch priamok. Smernica priamky od prvej pozície kamery po pozíciu objektu je určená ako p_1 , a zároveň smernica priamky od druhej pozície kamery po pozíciu objektu je určená ako p_2 . Avšak pre čitateľnosť bude od teraz platiť:

$$m_1 = p_1 \quad (2.40)$$

$$m_2 = p_2 \quad (2.41)$$

Bez odvádzania zavedieme funkciu $\varphi(a, b)$, ktorá pre dve priamky určené smernicami a a b určí uhol, ktorý tieto dve priamky zvierajú. Funkcia $\varphi(a, b)$ je definovaná ako:

$$\varphi(a, b) = \tan^{-1} \left(\frac{a - b}{1 + ab} \right) \quad (2.42)$$

Cieľovú hodnotu x_t určíme nasledovne:

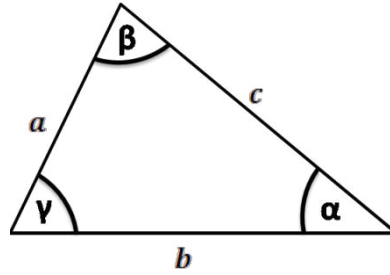
$$x_t = x_1 + \sin(\pi - \varphi(m_0, m_2)) \left(\frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{\sin(\varphi(m_0, m_2) - \varphi(m_0, m_1))} \right) \cos(m_1) \quad (2.43)$$

Cieľovú hodnotu y_t vypočítame nasledovne:

$$y_t = y_1 + \sin(\pi - \varphi(m_0, m_2)) \left(\frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{\sin(\varphi(m_0, m_2) - \varphi(m_0, m_1))} \right) \sin(m_1) \quad (2.44)$$

Nasleduje podrobný popis odvodenia jednotlivých vzťahov. Ako už bolo vopred avizované, základom je použitie sínusovej vety, ktorú uvádzame pre pripomenutie, nakoľko je kľúčová. Pre všeobecný trojuholník platí:

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma} \quad (2.45)$$



Obrázok 2-18 Trojuholník s uvedenými stranami a uhlami pre demonštráciu sínusovej vety.

Ak uhol α je uhol medzi priamkou spájajúcou prvú pozíciu kamery a pozíciu objektu (strana b), a priamkou, ktorá spája prvú a druhú pozíciu kamery (strana c), ktorý je určený ako $\varphi(m_0, m_1)$, a uhol β je rovnako určený pomocou $\varphi(m_0, m_2)$, avšak ako doplnok do hodnoty π , čiže $\beta = \pi - \varphi(m_0, m_2)$, potom uhol γ prináleží skutočnej polohe objektu. Pre uhol γ platí nasledovné odvodenie:

$$\gamma = \pi - (\alpha + \beta) \quad (2.46)$$

$$\gamma = \pi - (\varphi(m_0, m_1) + \pi - \varphi(m_0, m_2)) \quad (2.47)$$

$$\gamma = \varphi(m_0, m_2) - \varphi(m_0, m_1) \quad (2.48)$$

Ak poznáme všetky potrebné uhly v trojuholníku, môžeme určiť vzdialenosť Δd o ktorú sa vozidlo pohlo medzi danými dvomi pozíciami. Jedná sa o triviálny výpočet Euklidovskej vzdialenosti dvoch bodov v priestore. Pre Δd platí vzťah:

$$\Delta d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2.49)$$

Úlohou je teraz vypočítať, o akú vzdialenosť je potrebné sa posunúť v smere AP na prvom obrázku od pozície kamery na prvej snímke. Inak povedané, vzhľadom na trojuholník vyššie, poznáme všetky uhly, rovnako aj stranu $c = \Delta d$, a je potrebné určiť dĺžku strany b , keďže to je vzdialenosť objektu od kamery na prvom snímku. Ak získame túto vzdialenosť, vieme jednoducho urobiť priestorovú transformáciu bodu (x_1, y_1) o daný uhol a vzdialenosť.

Zo sínusovej vety vyššie (vzhľadom na značenie v trojuholníku na obrázku) môžeme zapísať takýto vzťah:

$$\frac{b}{\sin \beta} = \frac{c}{\sin \gamma} \quad (2.50)$$

Ak vzdialenosť, o ktorú je potrebné transformovať pozíciu kamery na prvom snímku označíme ako Δs , potom po dosadení zvyšných rovností platí:

$$\frac{\Delta s}{\sin(\pi - \varphi(m_0, m_2))} = \frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{\sin(\varphi(m_0, m_2) - \varphi(m_0, m_1))} \quad (2.51)$$

$$\Delta s = \sin(\pi - \varphi(m_0, m_2)) \left(\frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{\sin(\varphi(m_0, m_2) - \varphi(m_0, m_1))} \right) \quad (2.52)$$

Prichádza na rad spomínaná transformácia daného bodu o nejakú vzdialenosť v danom uhle. V našom prípade sa jedná o transformáciu bodu (x_1, y_1) o vzdialenosť Δs v uhle m_1 . Pre finálny bod (x_t, y_t) platí:

$$x_t = x_1 + \Delta s \cdot \cos m_1 \quad (2.53)$$

$$y_t = y_1 + \Delta s \cdot \sin m_1 \quad (2.54)$$

Po dosadení vyššie odvodených rovností získavame výslednú formu rovníc prezentovaných na začiatku.

$$x_t = x_1 + \sin(\pi - \varphi(m_0, m_2)) \left(\frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{\sin(\varphi(m_0, m_2) - \varphi(m_0, m_1))} \right) \cos(m_1) \quad (2.55)$$

$$y_t = y_1 + \sin(\pi - \varphi(m_0, m_2)) \left(\frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{\sin(\varphi(m_0, m_2) - \varphi(m_0, m_1))} \right) \sin(m_1) \quad (2.56)$$

2.7 Vylepšovanie odhadov jednotlivých metód

Metódy v oblasti PV často operujú s istou dávkou neurčitosti, ktorá je prítomná aj pri našich odhadoch. Ako bolo spomenuté, metóda je použitá na čo najväčší počet snímkov,

ktoré obsahujú konkrétny objekt záujmu, aby sme získali čo najväčší počet odhadov, a tieto neskôr vylepšili.

V dôsledku rôznych druhov chýb a iných negatívnych vplyvov samotná metóda bude vykazovať nejaké rozdelenie pravdepodobnosti distribúcie bodov (odhadov) v priestore, ktoré by malo byť intenzívne v okolí skutočnej pozície objektu, ktorý je snaha lokalizovať, a riedke v iných oblastiach priestoru. Pokiaľ takýto predpoklad nie je splnený tak je vysoká šanca, že samotná metóda je nekorektná, ak by sa potenciálne korektné odhady pozície štatisticky správali ako šum.

Šum je v našom probléme schopný výsledky ovplyvniť do veľkej miery. Na základe tohto je potrebné redukovať vplyv šumu v maximálnej možnej miere. Ak abstrahujeme od samotného vplyvu šumu napríklad nejakého GPS zariadenia, stále sú faktory, ktoré výpočet môžu negatívne ovplyvniť.

Rozlíšenie obrázka má vysoký vplyv na možnú chybovosť, pretože jeden pixel, ktorý predstavuje elementárny prvok obrázka, v konečnom dôsledku obsahuje úsek možných pozícií dopravného značenia, pričom všetky vyhodnotíme ako na jednej pozícií. V istých situáciách sú prítomné body, pozícia ktorých je veľmi vzdialená od hľadanej. V takýchto prípadoch sa jedná o tzv. *outliers* [17, s. 96].

Pod pojmom *inliers* (opak k *outliers*) budeme rozumieť body, ktoré sú relatívne blízko objektu (dopravného značenia), a teda ich môžeme považovať za potenciálnych kandidátov na polohu skúmaného objektu, a spomínaným pojmom *outliers* budeme označovať nepresné, chybové body.

Nakoľko metóda má produkovať väčšinu bodov v blízkosti s „relatívne malou“ odchýlkou vzhľadom na presnosť a možnosti metódy, *outliers* budú vykazovať často vysokú odchýlku od väčšiny bodov. Naším cieľom je tieto *outliers* korektne detegovať, odstrániť, a polohu objektu odhadovať len na základe bodov, ktoré podľa našej metriky do skupiny *outliers* nepatria. *Outliers* nesmú byť nejakým spôsobom organizované, a teda ich detekcia by mala byť realizovateľná.

V našom prípade bolo cieľom tieto *outliers* systematicky odstrániť a docieľiť tak zlepšenie odhadu danej pozície. Ako sme k tomuto problému pristupovali opäť zahŕňa pár rozličných spôsobov, ktoré detailne popíšeme v nasledujúcom texte.

Je nutné uviesť, že vylepšenia je možné realizovať len vďaka tomu, že často budeme disponovať väčším počtom odhadov polohy rovnakého objektu. Ak uvažujeme príklad videozáznamu pri vzorkovacej frekvencii 25 snímok za sekundu a objekt je viditeľný po dobu 2 sekúnd, toto rezultuje v 50 snímok, na ktorých sa daný objekt nachádza, a následne je možné uskutočniť $\binom{50}{2} = 1225$ meraní v najlepšom možnom prípade. Je samozrejmosťou, že veľa takýchto prípadov bude zahŕňať situácie, v ktorých výpočet nie je odporúčaný a budú teda odfiltrované, ale i napriek tomu získame relatívne vysoký počet párov snímok, ktoré môžeme pre našu metódu použiť. Takéto odhady však budú obsahovať *outliers* v prípadoch, ktoré budú blízko situáciám, ktoré považujeme za problémové, a tieto body je potrebné z výsledkov systematicky odstrániť.

2.7.1 Globálny centroid pre odhad pozície

V našich experimentoch sme realizovali aj naivný, triviálny postup pre odhad pozície, ktorý zahŕňal výpočet globálneho centroidu pre všetky získané odhady pozície a tento bol následne prehlásený ako vylepšený odhad.

Jednoduchosť tejto metódy so sebou prináša veľa nevýhod, a je pre praktické účely nepoužiteľná ako samostatná metóda, a pri testovaní sme tento prístup použili len ako referenciu pre porovnanie s nejakou inou, sofistikovanejšou.

Je nutné však podotknúť, že nápad vypočítať centroid z danej množiny bodov je kľúčový pre vylepšenie odhadu pozície, ale je potrebné ho doplniť ešte o iné postupy, ktoré jeho nevýhody značne zredukujú. Jeden takýto uvedieme v nasledujúcej kapitole.

Ako je zrejmé, centroid predstavuje „priemerný bod“ z množiny bodov, a priemer ako taký je veľmi citlivý na extrémne hodnoty, čo je hlavný dôvod, prečo tento prístup má vysokú chybovosť. Avšak pri korektnom odstránení spomínaných *outliers*, potom už je priemer použiteľný.

2.7.2 Zhľukovanie pre odstraňovanie šumu

Problém použitia len správnych kandidátov pre odhad reálnej pozície sme sa rozhodli riešiť heuristicky metódou, ktorá ideologicky inklinuje k mnohým algoritmom z oblasti zvanej zhľukovanie (angl. *clustering*) [17, s. 385], ktorá v skratke poskytuje aparát pre rozdelenie množiny bodov do zhľukov na základe konkrétneho súboru vlastností meraných špecifickou metrikou. Túto metódu sme vymysleli priamo pre naše účely, a nedá sa priamo hovoriť o zhľukovaní, pretože u nás nie sú jednotlivé zhľuky disjunktné. Jedná sa len o principiálne veľmi podobný postup.

Uvažujme množinu bodov, ktorú vyprodukovala naša metóda z niekoľkých párov obrázkov, na základe ktorých boli vykonané výpočty. Úlohou je teraz určiť taký zhhluk bodov, v ktorom je smerodajná odchýlka všetkých vzájomných vzdialeností bodov minimálna, a teda je vysoká šanca (na základe predpokladu vyššie spomenutého ohľadom *inliers* a *outliers*), že sa jedná o zhhluk bodov, ktoré reprezentujú možných kandidátov na skutočnú pozíciu.

Po určení veľkosti zhhluku (pravidlá a výpočet popíšeme neskôr) vyberieme pre každý bod niekoľko (počet ekvivalentný veľkosti zhhluku určenej pred začiatkom algoritmu) k nemu najbližších bodov, a vypočítame smerodajnú odchýlku vzdialeností bodov v rámci tohto zhhluku. Tento postup aplikujeme pre všetky body. Je nutné opäť zdôrazniť, že jednotlivé zhhluky nie sú disjunktné, čiže počet zhhlukov je rovnaký počtu bodov. Zhhluk, ktorého smerodajná odchýlka vzdialeností bodov je minimálna spomedzi všetkých sa stane kandidátom pre odhad pozície objektu. Túto pozíciu určíme ako centroid bodov vybraného zhhluku. Teraz postup popíšeme presnejšie.

Experimentálne sme určili, že počet vstupných bodov musí byť **aspoň 5** pre vhodné fungovanie metódy. Číslo 5 reprezentuje počet odhadov pozície, čiže 5 párov snímok, na základe ktorých boli tieto odhady vykonané. V texte nižšie budeme túto podmienku považovať za splnenú, teda predpokladáme $n \geq 5$. Algoritmus by pracoval korektne aj s menším počtom bodov avšak degradoval by na klasický výpočet centroidu všetkých bodov a nepriniesol by žiadnu výhodu.

Nech S je cieľová veľkosť zhhluku, hodnota ktorej je určená nižšie:

$$S = \max(3, \lceil \sqrt{n} \rceil) \quad (2.57)$$

Pre takýto vzťah sme sa rozhodli z nasledovných dôvodov. Uvedomujeme si náchylnosť výpočtu výsledného odhadu na *outliers*, a na základe tohto je vhodné množinu elementárnych odhadov z ktorých bude výsledný odhad skonštruovaný udržať s menším počtom prvkov. Prichádza tu aj výhoda z výpočtového hľadiska. Avšak na druhej strane je potrebné počet prvkov tejto množiny ohraničiť aj zdola, a číslo 3 nám empiricky pripadalo ako vhodná hranica pre takýto druh výpočtu.

Pri väčšom počte bodov ako je 9 je použitá horná hranica odmocniny z tohto počtu. Nápadom bolo určiť veľkosť zhhluku transformáciou, ktorá pri malých počtoch rastie viacej ako pri veľkých, nakoľko od istej hranice už je šanca, že sme zachytili dostatočne veľkú množinu bodov v okolí „korektnej“ pozície objektu. Rozhodli sme sa pre odmocninu, ale

aj transformácia pomocou logaritmu alebo inej funkcie s podobnými vlastnosťami by bola postačujúca. Vo veľa prípadoch je počet dostupných snímok, na ktorých sa nachádza daný objekt, vysoký už po pár sekundách snímania, a produkovanie párov rastie kvadraticky, čiže často bude algoritmus mať k dispozícii stovky, niekedy aj tisíce odhadov polohy. Experimentálne sme overili lineárne rastúcu veľkosť zhluku akými boli $\frac{n}{2}$, $\frac{n}{3}$ a podobne, a výpočet nielenže bol časovo náročnejší, výsledná pozícia bola odhadnutá horšie (aj keď v priemere len o 10% vzhľadom na nami nájdený, najlepší odhad). Použitie $\lceil \sqrt{n} \rceil$ namiesto $\lfloor \sqrt{n} \rfloor$ opäť produkuje približne o 8% lepšie výsledky, to nás privádza k názoru, že existuje nejaká teoreticky optimálne stanovená veľkosť zhluku. Aby sme spresnili význam percentuálnych tvrdení, rozdiely boli len na úrovni centimetrov, čiže všetky odhady boli veľmi uspokojivé bez ohľadu na funkciu. Na základe tohto zotrvávame pri nelineárnej funkcii, ale pripúšťame možnosť hlbšieho experimentovania v budúcnosti.

Po určení veľkosti zhluku je následne potrebné vytvoriť množinu zhukov, ktoré obsahujú po S prvkoch. Túto množinu vytvoríme na základe Euklidovskej vzdialenosti dvoch bodov nasledovne:

$$C = \{\varphi(\{(p_j, \|p_j - p_i\|) \mid j = 1 \dots n, j \neq i\}, S) \mid i = 1 \dots n\} \quad (2.58)$$

Kde funkcia φ berie ako prvý parameter množinu bodov spoločne so vzdialenosťami k aktuálne spracovávanému bodu a vráti S najlepších bodov. Tieto body budeme považovať za jednotlivé zhuky.

Je evidentné, že $|C| = n$, kde $|C|$ predstavuje počet prvkov v množine C . Pre každý bod musí pripadať nejaký zhuk bodov rovnakej veľkosti (S).

Spomedzi vytvorených zhukov je nutné nájsť taký, v rámci ktorého smerodajná odchýlka všetkých vzájomných vzdialeností bodov je minimálna.

Index daného zhluku (k) z množiny zhukov C získame nasledovným spôsobom:

$$k = \text{solvemin}\{\omega(C_i) \mid i = 1 \dots |C|\} \quad (2.59)$$

Kde funkcia ω vypočíta smerodajnú odchýlku všetkých vzájomných vzdialeností bodov daného zhluku, a je definovaná nasledovne:

$$\omega(c) = \frac{2}{S(S-1)-2} \sum_{i=1}^{S-1} \sum_{j=i+1}^S (\|c_i - c_j\| - \bar{d})^2 \quad (2.60)$$

Kde $\bar{d}$ reprezentuje priemernú vzdialenosť všetkých bodov, pre ktorú platí:

$$\bar{d} = \frac{2}{S(S-1)} \sum_{i=1}^{S-1} \sum_{j=i+1}^S \|c_i - c_j\| \quad (2.61)$$

Cieľom je v prvom rade určiť hodnotu $\bar{d}$ a to výpočtom sumy vzájomných vzdialeností všetkých možných dvojíc z S bodov, čomu korešpondujú dve sumy vo výpočte. Tento súčet je následne podelený počtom všetkých možných dvojíc bodov z S bodov, ktorý je rovný $\binom{S}{2} = \frac{S(S-1)}{2}$, a jeho prevrátená hodnota je prítomná vo vzorci ako $\frac{2}{S(S-1)}$. V prípade smerodajnej odchýlky platí pre výraz $\frac{2}{S(S-1)-2}$ rovnaká úvaha, pričom odčítanie čísla 2 dôsledkom delenia počtom $n - 1$, ktorý sa používa pre nevychýlený odhad vzorky smerodajnej odchýlky.

Po získaní zhľuku, v rámci ktorého je priemerná vzdialenosť minimálna môžeme určiť jeho centroid, čím získame vylepšený odhad reálnej pozície objektu.

Pozíciu $\bar{x}_t$ získame ako:

$$\bar{x}_t = \frac{1}{S} \sum_{i=1}^S C_{ki0} \quad (2.62)$$

Pozíciu $\bar{y}_t$ získame ako:

$$\bar{y}_t = \frac{1}{S} \sum_{i=1}^S C_{ki1} \quad (2.63)$$

V tomto kroku sme učinili spresnenie odhadu pozície objektu transformáciou bodov x_t a y_t na body $\bar{x}_t$ a $\bar{y}_t$ procesom vyššie popísaným. Tieto body sú výstupom z tejto metódy ako odhad pozície objektu.

2.8 Potenciál pre vylepšenie trasovania objektov

V tejto krátkej, teoretickej podkapitole uvedieme zatiaľ neoverené hypotézy o novej využiteľnosti našich výsledkov aj pre trasovanie objektu na základe jeho BB. Triviálna forma trasovania je prítomná aj v tejto práci v prípade RD, nakoľko anotácia objektu poskytuje len BB a oficiálnu kategóriu dopravného značenia (ktorá nie je unikátna), ale pre naše účely potrebujeme určiť, či sa jedná o rovnaký objekt.

Existuje mnoho algoritmov pre trasovanie objektov, avšak vo svojej podstate sa jedná o vytvorenie sledu (v literatúre označovaný aj ako *track*) niekoľkých BB a na základe ich vlastností je vybudovaný odhad, že patria identickému objektu.

Situácia však nie je vždy ideálna a pre demonštráciu, kde by sa dali aplikovať naše výsledky uvažujme prípad, kedy sa na jednej snímke nachádza viacej podobných alebo rovnakých objektov (napríklad dopravných značení), ktoré sú relatívne blízko seba. V takýchto prípadoch môže BB pri trasovaní doslova „preskočiť“ z jedného objektu na druhý napríklad v dôsledku pohybu vozidla v ostrej zákrute a pod.

Našou úvahou je využitie odhadovania pozície objektu a trasovanie objektu vylepšiť. Postup nemusí poskytovať funkcionality opravenia samotného trasovania, ale len detekciu chýb v dátach, ktoré by neskôr mohli byť opravené buď manuálne alebo automaticky. Na základe tohto sa jedná len o kontrolný mechanizmus.

Dramatická zmena BB spôsobí v našich metódach odchýlku od skutočnej hodnoty, pretože spomínaný „lúč“ bude vychýlený v odlišnom uhle. Týmto spôsobom by bolo možné sekvenciu poskytnutých odhadov pre jednotlivé BB analyzovať odhadovaním polohy a ak dôjde k vysokej odchýlke, takýto *track* označiť za rizikový.

Vzhľadom na to, že je nutné definovať odchýlku kvantitatívne, tak táto úvaha nás privádza k ďalšiemu dôležitému aspektu tejto práce, a to je možnosť kalkulácie úrovne dôveryhodnosti v náš odhad. Vynára sa tú možnosť použiť viacero parametrov, akými sú rozdiel v uhloch pod akými sú objekty viditeľné, ich veľkosti BB a pod. Za pomoci týchto hodnôt by sme mohli ideálne vyjadriť koeficient z intervalu $\langle 0, 1 \rangle$, ktorý by reprezentoval úroveň dôveryhodnosti.

Ďalšou rizikovou situáciou je výpadok BB pre konkrétny objekt. Chýbajúce záznamy v sekvencii BB je možné riešiť triviálnym spôsobom založeným na interpolácii, ktorý bol použitý aj v našej práci. V takýchto prípadoch by použitie našej metódy nemuselo vôbec prispieť k zlepšeniu výsledkov, avšak implementačná náročnosť by bola podstatne zvýšená.

Za úvahu stoja situácie, kedy výpadok BB je dlhodobejší, kedy triviálne algoritmy trasovania nemôžu ani s vysokou dávkou dôveryhodnosti povedať, že sa jedná o rovnaký objekt. Problém by bolo možné riešiť použitím našich metód priamo a aj inverzne. Pod priamym použitím myslíme realizáciu „klasického“ výpočtu odhadovania pozície, pričom pod inverzným myslíme výpočet, kde by sa mal BB daného objektu v rámci snímky nachádzať, ak sa jedná o objekt na nejakej konkrétnej pozícii. V takomto prípade by bolo možné aj dlhodobejšie výpadky aspoň čiastočne zrekonštruovať.

3 Výsledky práce a diskusia

V tejto kapitole uvedieme výsledky našej práce na všetkých typoch *datasetov*, ktoré sme predstavili v prechádzajúcej kapitole. Pri každom type poskytneme štatistické ukazovatele, na základe ktorých sme došli k našim záverom ohľadom presnosti a použiteľnosti v jednotlivých situáciách.

Najskôr uvedieme krátky komentár k softvérovej časti procesu testovania, a následne všeobecný výpočet metriky presnosti, ktorý bol použitý pre všetky typy *datasetov*. Neskôr pomocou tabuliek poskytneme sumárne výsledky testovania s dodatočným komentárom. Pri špecifických situáciách bude prítomná aj grafická ilustrácia situácie pre ľahšie pochopenie a skúmanie konkrétneho prípadu testovania.

3.1 Implementácia testovacej aplikácie

Pre účely testovania bola vytvorená aplikácia pomocou programovacieho jazyka C# v kombinácii s WPF (Windows Presentation Foundation) pre realizáciu GUI (Graphical User Interface). Táto aplikácia nie je cieľovým produktom, preto je komentáru venovaná len malá časť. Program slúžil primárne na uľahčenie testovania, vizualizácie situácií a experimentovania.

Objektovo orientovaný prístup s použitím niekoľko štandardných návrhových vzorov a implementácie vlastného mechanizmu pre architektonický štýl zvaný Model-View-ViewModel (MVVM) nám umožnil zabezpečiť vysokú znovu použiteľnosť a rozširiteľnosť, takže pridávanie nových *datasetov*, metód kalkulácie prípadne metód zlepšovania odhadov je triviálne.

Okrem iného, vo veľkej miere bol použitý programovací jazyk Python na konverziu *datasetov* z ich „surových“ formátov do „štandardného“ pre možné spracovanie aplikáciou predstavenou vyššie. Štatistické analýzy boli rovnako realizované za pomoci tohto jazyka.

3.2 Metrika presnosti metód

Chybu E vypočítame ako Euklidovskú vzdialenosť medzi odhadnutým bodom v 2D priestore a skutočnou polohou objektu, ktorú naše syntetické dáta poskytujú. Nech R_x a R_y reprezentujú skutočnú polohu objektu, a body S_x a S_y polohu odhadnutú. Platí:

$$E = \sqrt{(R_x - S_x)^2 + (R_y - S_y)^2} \quad (3.1)$$

Chyba E teda reprezentuje globálnu metriku našej metódy, na základe ktorej sme presnosť a použiteľnosť v jednotlivých testovaných situáciách posudzovali.

Pre minimalizáciu chybovosti bolo potrebné odstraňovať nevhodné situácie, ktoré by matematicky predstavovali problémy. Tieto sú však špecifické pre jednotlivé metódy. V predchádzajúcej kapitole sa nachádza popis metód odstraňovania *outliers*. Výstupom takéhoto procesu je finálny odhad polohy objektu, ktorý je neskôr pomocou Euklidovskej vzdialenosti porovnaný s „korektnou“ polohou objektu. Avšak táto metrika je použitá aj pre výpočet chybovosti v prípade jednotlivých dvojíc snímok, nie len pre finálny odhad.

3.3 Konverzia súradníc zemepisnej šírky a dĺžky do Kartézskoho súradnicového systému

V prípade RD, v ktorom sú pozície objektov a kamery získane GPS zariadením, je potrebné pre naše výpočty súradnice zemepisnej šírky a výšky previesť na klasické x , y a z body Kartézskoho súradnicového systému, s ktorým už naša implementácia metód dokáže pracovať.

Nech α_{lat} je uhol zemepisnej šírky a β_{lon} je uhol zemepisnej dĺžky a R nech je polomer zemegule približne rovný 6 378 000 metrov. Prevod na súradnice x , y a z vykonáme nasledovne:

$$x = R \cos(\alpha_{lat}) \cos(\beta_{lon}) \quad (3.2)$$

$$y = R \cos(\alpha_{lat}) \sin(\beta_{lon}) \quad (3.3)$$

$$z = R \sin(\alpha_{lat}) \quad (3.4)$$

Všetky spomenuté vzťahy ako aj výpočet metriky sú na tieto súradnice aplikované rovnako ako v prípade SD. Pripúšťame možnosť dodatočnej chybovosti pri konverzii a následných výpočtoch s reálnou aritmetikou počítača v prípade trigonometrických funkcií.

Spätný prevod je realizovaný ako:

$$\alpha_{lat} = \sin^{-1}\left(\frac{z}{R}\right) \quad (3.5)$$

$$\beta_{lon} = \text{atan2}(y, x) \quad (3.6)$$

V rovnici 3.6 sme použili funkciu $\text{atan2}(y, x)$ [21], ktorá je definovaná ako uhol medzi vektorom (x, y) a x -ovou osou. Funkcia výpočet podstatne zjednodušuje nakoľko korektne určuje veľa podmienok, ktoré pri uhloch môžu nastať vzhľadom na rôzne kvadranty. Táto funkcia je dostupná v štandardných knižniciach programovacích jazykov.

3.4 Testovanie, spracovanie a prezentovanie výsledkov

Získavanie výsledkov a následne spracovanie pre jednotlivé *datasety* bude pozostávať z niekoľkých krokov. Najskôr bude nad daným *datasetom* spustené dávkové testovanie pre všetky dostupné objekty, ktoré sa v *datasete* nachádzajú. Tento proces spomedzi všetkých možných dvojíc snímok, na ktorých je konkrétny objekt viditeľný, použije na výpočet tie, ktoré rešpektujú nami stanovené obmedzenia (vlastnosti uhlov a pod.). Jednotlivé výsledky budú zaznamenané, a nad týmito jednotlivými odhadmi budú vykonané štatistické analýzy.

Predmetom záujmu z hľadiska štatistiky budú minimálna chyba a maximálna chyba, aby sme získali prehľad o celkovom rozsahu chybovosti. Následne je dôležitý priemer, avšak ten musí byť patrične doplnený aj smerodajnou odchýlkou, nakoľko sám o sebe by nemal moc vysokú výpovednú hodnotu. Vzhľadom na to, že pre ľudí je ťažšie interpretovať dvojicu priemer a smerodajná odchýlka a získať prehľad o tom, čo je možné čakať, tak poskytneme aj medián. Týmto piatimi hodnotami bude popísaný výsledok každého *datasetu*, ktoré budú neskôr globálne porovnané a vyhodnotené.

Je potrebné zdôrazniť, že v tejto analýze nehrá rolu odstraňovanie *outliers*, pretože chceme získať prehľad o jednotlivých, “surových“ odhadoch.

Ďalším krokom bude aplikácia oboch metód špecificky vytvorených pre určenie výsledného odhadu polohy (výpočet pomocou globálneho centroidu ako aj pomocou zhľukovania). Tieto merania zahrnieme do separátnej tabuľky pre jednotlivé *datasety*, kde porovnáme aj účinnosť metódy zhľukovania voči primitívnej kalkulácii centroidu všetkých bodov. Výsledné odhady naprieč všetkými SD, ktorých bude rovnaký počet ako je celkový počet skúmaných objektov, budú podrobené identickej štatistickej analýze predstavenej vyššie, ktorá bude rovnako pozostávať z 5 ukazovateľov. Minimálna a maximálna chyba najlepšieho odhadu od skutočnej pozície objektu, spoločne s priemerom, smerodajnou odchýlkou a mediánom.

V prípade SD budeme skúmať aj aký vplyv má šum na výsledky procesu. Šumom budeme rozumieť náhodné modifikovanie polohy kamery, alebo azimutu kamery, prípadne BB objektu na základe nejakých parametrov, ktoré by odpovedali realite (normálne rozdelenie pravdepodobnosti šumu a pod.). Typ šumu a bližší popis uvedieme spolu s výsledkami neskôr.

Ako posledný fakt, ktorý je treba podotknúť je, že tento proces bude identický pre obidve metódy odhadovania pozície objektu, pretože jedným z cieľom je porovnať aj presnosť jednotlivých metód. Pre porovnanie metód uvedieme aj vizualizáciu pomocou tzv. *scatter plot*, ktorá bude ilustrovať distribúciu odhadov pozície v priestore. Vopred avizujeme, že výsledky vyzerajú dramaticky nepresné len na pohľad, nakoľko je dôležité si uvedomiť mieru škálovania.

3.4.1 Výsledky testovania pre syntetické *datasets*

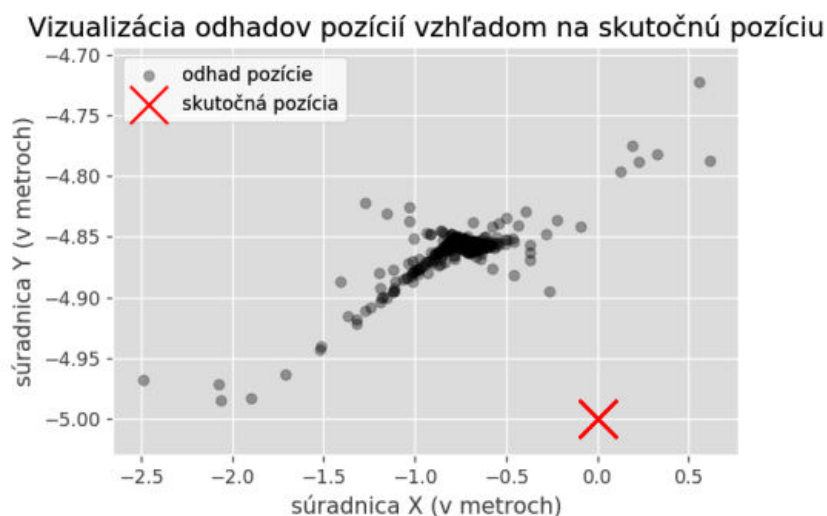
Dataset flat

Metóda	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	0.185	2.488	0.789	0.216	0.761
Trojuholník	0.185	2.488	0.790	0.216	0.762

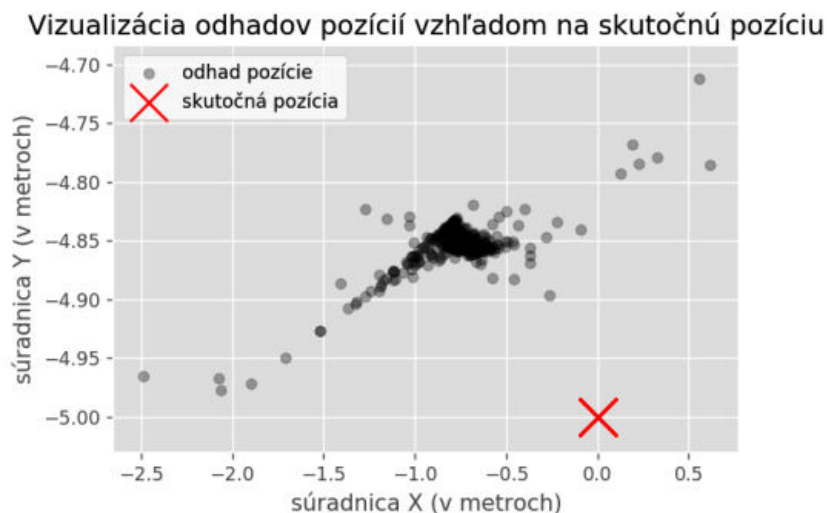
Tabuľka 4 Štatistika chybovosti (v metroch) jednotlivých odhadov pre *dataset flat*.

Metóda lokalizácie	Metóda vylepšenia	Chyba
Priamky	Primitívny centroid	0.777
Trojuholník	Primitívny centroid	0.779
Priamky	Zhlukovanie	0.763
Trojuholník	Zhlukovanie	0.724

Tabuľka 5 Štatistika chybovosti (v metroch) výsledného odhadu pre *dataset flat*.



3-1 Vizualizácia odhadov - *dataset: flat*, metóda: priamky.

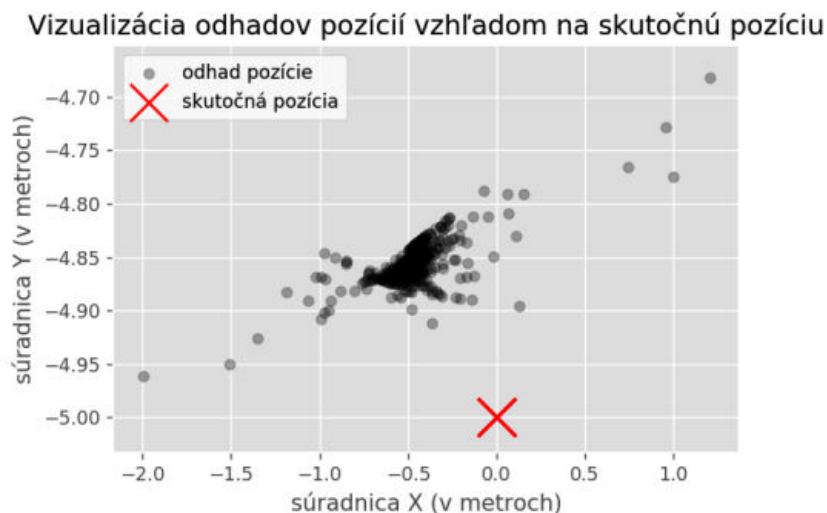
3-2 Vizualizácia odhadov - *dataset: flat*, metóda: trojuholník.**Dataset hill**

Metóda	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	0.108	2.125	0.646	0.173	0.640
Trojuholník	0.151	1.996	0.519	0.174	0.499

Tabuľka 6 Štatistika chybovosti (v metroch) jednotlivých odhadov pre *dataset hill*.

Metóda lokalizácie	Metóda vylepšenia	Chyba
Priamky	Primitívny centroid	0.629
Trojuholník	Primitívny centroid	0.555
Priamky	Zhlukovanie	0.600
Trojuholník	Zhlukovanie	0.585

Tabuľka 7 Štatistika chybovosti (v metroch) výsledného odhadu pre *dataset hill*.3-3 Vizualizácia odhadov - *dataset: hill*, metóda: priamky.

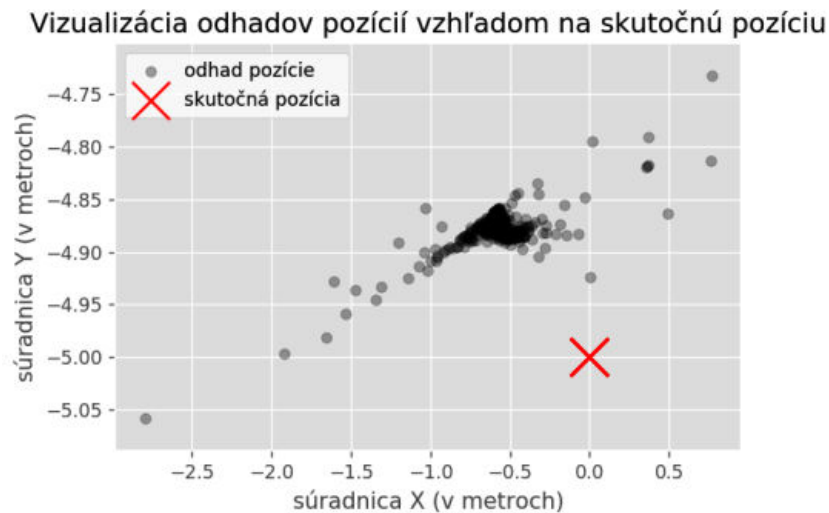
3-4 Vizualizácia odhadov - *dataset: hill*, metóda: trojuholník.**Dataset up**

Metóda	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	0.083	2.828	0.637	0.213	0.614
Trojuholník	0.076	2.794	0.605	0.211	0.580

Tabuľka 8 Štatistika chybovosti (v metroch) jednotlivých odhadov pre *dataset up*.

Metóda lokalizácie	Metóda vylepšenia	Chyba
Priamky	Primitívny centroid	0.621
Trojuholník	Primitívny centroid	0.603
Priamky	Zhlukovanie	0.617
Trojuholník	Zhlukovanie	0.663

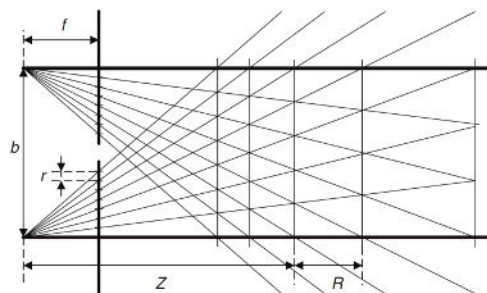
Tabuľka 9 Štatistika chybovosti (v metroch) výsledného odhadu pre *dataset up*.3-5 Vizualizácia odhadov - *dataset: up*, metóda: priamky.



3-6 Vizualizácia odhadov - dataset: *up*, metóda: trojuholník.

Pri SD typu *flat*, *hill* a *up* je pri vizualizácii pomocou *scatter plot* možné vidieť podobný trend odhadov. Metóda je konzistentne vychýlená voči nejakému bodu v priestore, okolo ktorého sa koncentruje väčšina odhadov. Niekoľko krát spomínané *outliers* sú zreteľne viditeľné, čím opodstatňujeme naše snahy k ich odstraňovaniu.

Dôvodov pre tento trend odchýlky je niekoľko, my sa prikláňame k hypotéze o rozlíšení obrázka a pozícií objektu. Rozlíšenie má priamy vplyv na to, akú plochu reálneho sveta pokrýva jeden pixel, a na základe tohto objekty, ktorých BB pripadá na takúto oblasť budú označené ako veľmi blízke. Nakoľko je pravdepodobnejšie získať polohu vychýlenú ako priamu, pretože počet „nesprávnych“ odhadov v rámci nejakého regiónu, v ktorom sú všetky objekty považované za veľmi blízke, skrátka prevyšuje počet „správnych“ odhadov.



3-7 Ukážka nerovnomerného rozloženia bodov pripadajúcich na 1 pixel [1, s. 3].

Ak analyzujeme súradnice týchto bodov vzhľadom na súradnicový systém SD a smer pohybu vozidla tak vidíme, že odhady sú „vľavo“ od skutočnej polohy. Toto môže byť zapríčinené tým, že dopravné značenie je umiestnené na pravej strane vozovky, kde

opäť distribúcia priestoru, ktorý je pokrytý jedným pixlom je redšia ako pri pixloch pokrývajúcich oblasť pred vozidlom. Ukážka na obrázku 3-7.

Okrem iného, priestorová odchýlka na všetky strany je vysvetlená typom cesty. Ak auto zabáča do strán, tak odhady sú rozptýlené do všetkých strán.

Dataset straight-left

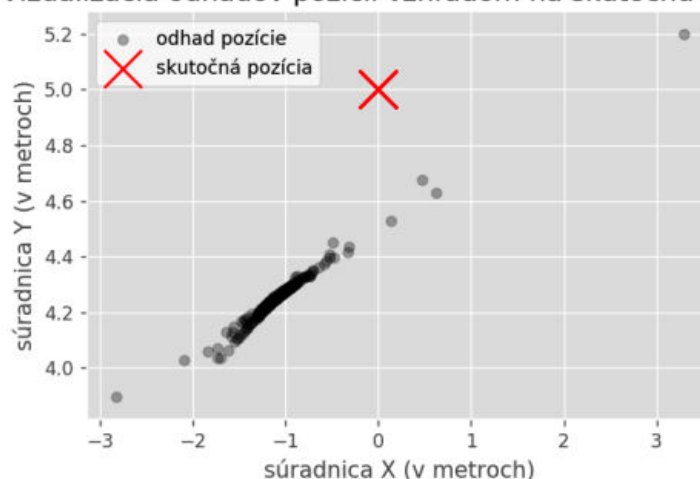
Metóda	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	0.493	3.296	1.364	0.228	1.359
Trojuholník	0.401	3.283	1.329	0.233	1.324

Tabuľka 10 Štatistika chybovosti (v metroch) jednotlivých odhadov pre *dataset straight-left*.

Metóda lokalizácie	Metóda vylepšenia	Chyba
Priamky	Primitívny centroid	1.343
Trojuholník	Primitívny centroid	1.307
Priamky	Zhlukovanie	1.337
Trojuholník	Zhlukovanie	1.305

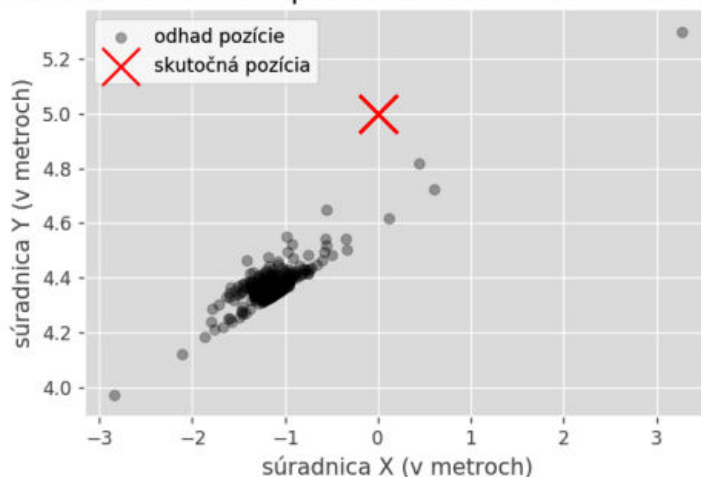
Tabuľka 11 Štatistika chybovosti (v metroch) výsledného odhadu pre *dataset straight-left*.

Vizualizácia odhadov pozícií vzhľadom na skutočnú pozíciu



3-8 Vizualizácia odhadov - *dataset: straight-left*, metóda: priamky.

Vizualizácia odhadov pozícií vzhľadom na skutočnú pozíciu

3-9 Vizualizácia odhadov - *dataset: straight-left*, metóda: trojuholník.

Dataset straight-left vykazuje najhoršie výsledky zo všetkých. Dokonca aj distribúcia bodov nie je podobná ako v prípade *datasetu straight-right* (nasleduje), čo by bolo očakávané. Možným dôvodom je umiestnenie dopravného značenia na ľavej strane, pričom auto sa pohybuje na pravej strane vozovky. Tento dodatočný rozdiel vo vzdialenosti môže spôsobovať takúto odchýlku, pretože uhol, pod ktorým je objekt viditeľný, je ostrejší. Dopravné značenie umiestnené na pravej strane v *datasete straight-right* vykazuje naopak výsledky výborné.

Dataset straight-right

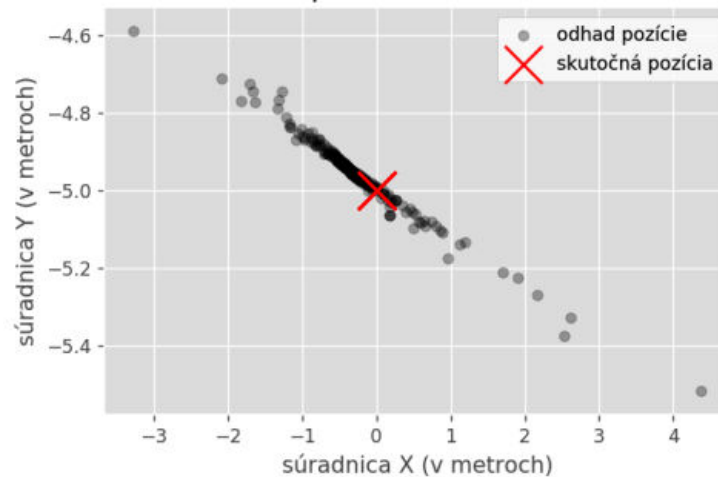
Metóda	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	0.004	4.408	0.458	0.405	0.389
Trojuholník	0.011	4.408	0.459	0.404	0.390

Tabuľka 12 Štatistika chybovosti (v metroch) jednotlivých odhadov pre *dataset straight-right*.

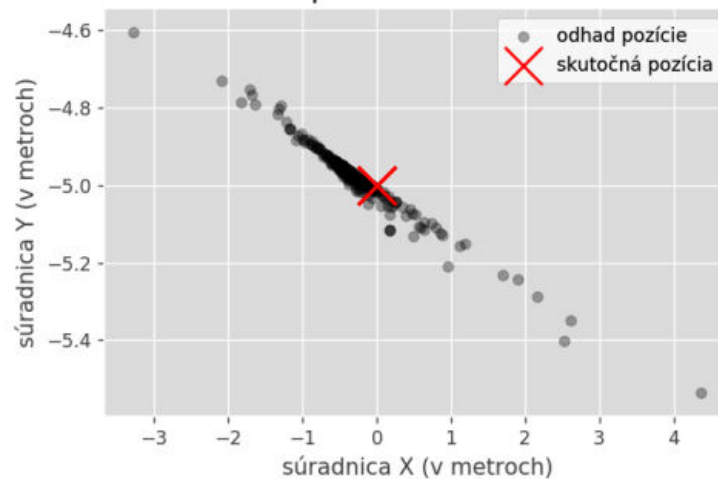
Metóda lokalizácie	Metóda vylepšenia	Chyba
Priamky	Primitívny centroid	0.318
Trojuholník	Primitívny centroid	0.319
Priamky	Zhlukovanie	0.356
Trojuholník	Zhlukovanie	0.353

Tabuľka 13 Štatistika chybovosti (v metroch) výsledného odhadu pre *dataset straight-right*.

Vizualizácia odhadov pozícií vzhľadom na skutočnú pozíciu

3-10 Vizualizácia odhadov - *dataset: straight-right*, metóda: priamky.

Vizualizácia odhadov pozícií vzhľadom na skutočnú pozíciu

3-11 Vizualizácia odhadov - *dataset: straight-right*, metóda: trojuholník.

Ak nahliadneme na odhady pre *dataset straight-right*, na ktorom je cesta rovná, tak vidíme, že odhady sú prevažne rozptýlené v smere jazdy vozidla, pred a za skutočnou polohou objektu. Opäť, v takomto prípade musí niekedy čisto zo štatistického hľadiska nastať, že pretneme skutočnú polohu objektu, nakoľko prechádzame po priamke.

3.4.2 Sumárne štatistiky výsledných odhadov pre syntetické *datasety*

V tejto krátkej časti uvedieme štatistiku chybovosti pre najlepšie odhady naprieč všetkými *datasetmi* vzhľadom na jednotlivé metódy kalkulácie pozície a aj ich zlepšovania.

Metóda lokalizácie	Metóda vylepšenia	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	Primitívny centroid	0.318	1.343	0.738	0.377	0.629

Trojuholník	Primitívny centroid	0.319	1.307	0.713	0.371	0.603
Priamky	Zhlukovanie	0.356	1.337	0.735	0.367	0.617
Trojuholník	Zhlukovanie	0.353	1.305	0.726	0.353	0.663

Tabuľka 14 Sumárna štatistika (v metroch) pre výsledné odhady pre syntetické *datasety*.

Metódy „priamky“ a „trojuholník“ vykazujú veľmi podobnú presnosť. Obidve sú prakticky použiteľné a všetky naše všeobecné tvrdenia platia pre obidve metódy. Tabuľky 4, 6, 8, 10 a 12 vykazujú, že obidve metódy majú problémy s podobným typom prípadov, čiže často hodnoty jednotlivých 5 štatistických ukazovateľov sú podobné, aj keď nie rovnaké. Toto nás privádza k názoru, že vlastnosti po aplikačnej stránke môžeme generalizovať na obidva prístupy.

Metódy vylepšenia odhadov nevykazujú vysoké odchýlky na SD. Je to spôsobené tým, že výpočty neobsahujú vysoké množstvo *outliers*, a teda aj samotný centroid môže byť použitý a odhad nebude príliš vychýlený. Avšak, ako bude možné vidieť v prípade RD, tam sú *outliers* častejšie a primitívny centroid v istých situáciách produkuje vyššiu chybovosť.

3.4.3 Výsledky testovania s umelým šumom pre syntetické *datasety*

Pre pridanie šumu boli všade použité nasledovné nastavenia. Šum pochádza z normálneho rozdelenia pravdepodobnosti. Stredná hodnota je vždy v hodnote, ktorú modifikujeme. Jediný parameter, ktorý bol špecifický je smerodajná odchýlka. Na jednotlivé súradnice kamery (*Pozícia k.*) v priestore bol aplikovaný šum so smerodajnou odchýlkou 0,25 metra. Azimut kamery (*Azimut k.*) bol modifikovaný rovnako šumom so smerodajnou odchýlkou 1°, a BB objektu s odchýlkou 10 pixlov (súradnice *X* a *Y* konkrétneho BB boli modifikované, šírka a výška ostali nezmenené).

Dataset flat

Metóda	Objekt šumu	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	Pozícia k.	0.057	53.648	4.835	7.840	2.080
Trojuholník	Pozícia k.	0.091	125.495	8.652	13.742	3.691
Priamky	Azimut k.	0.079	3341.935	19.616	162.621	3.976
Trojuholník	Azimut k.	0.172	22394.616	76.343	1088.374	4.426
Priamky	BB	0.102	1013.828	7.024	49.946	1.752
Trojuholník	BB	0.046	155.409	6.275	15.884	1.949

Tabuľka 15 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre *dataset flat*.

Objekt šumu	Metóda lokalizácie	Metóda vylepšenia	Chyba
-------------	--------------------	-------------------	-------

Pozícia k.	Priamky	Primitívny centroid	0.271
Pozícia k.	Trojuholník	Primitívny centroid	0.122
Pozícia k.	Priamky	Zhlukovanie	2.158
Pozícia k.	Trojuholník	Zhlukovanie	0.203
Azimut k.	Priamky	Primitívny centroid	7.339
Azimut k.	Trojuholník	Primitívny centroid	57.765
Azimut k.	Priamky	Zhlukovanie	1.001
Azimut k.	Trojuholník	Zhlukovanie	3.057
BB	Priamky	Primitívny centroid	1.944
BB	Trojuholník	Primitívny centroid	0.746
BB	Priamky	Zhlukovanie	1.038
BB	Trojuholník	Zhlukovanie	1.164

Tabuľka 16 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre *dataset flat*.

Dataset hill

Metóda	Objekt šumu	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	Pozícia k.	0.110	112.547	7.355	12.449	3.097
Trojuholník	Pozícia k.	0.028	164.529	6.412	13.111	2.804
Priamky	Azimut k.	0.087	1451.371	21.181	106.877	3.731
Trojuholník	Azimut k.	0.050	354.683	8.992	26.676	3.222
Priamky	BB	0.022	369.991	4.854	21.275	1.618
Trojuholník	BB	0.046	282.609	4.819	17.328	1.593

Tabuľka 17 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre *dataset hill*.

Objekt šumu	Metóda lokalizácie	Metóda vylepšenia	Chyba
Pozícia k.	Priamky	Primitívny centroid	0.275
Pozícia k.	Trojuholník	Primitívny centroid	1.576
Pozícia k.	Priamky	Zhlukovanie	0.333
Pozícia k.	Trojuholník	Zhlukovanie	1.422
Azimut k.	Priamky	Primitívny centroid	0.390
Azimut k.	Trojuholník	Primitívny centroid	7.921
Azimut k.	Priamky	Zhlukovanie	2.383
Azimut k.	Trojuholník	Zhlukovanie	1.077
BB	Priamky	Primitívny centroid	1.409
BB	Trojuholník	Primitívny centroid	0.476
BB	Priamky	Zhlukovanie	0.172
BB	Trojuholník	Zhlukovanie	0.444

Tabuľka 18 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre *dataset hill*.

Dataset up

Metóda	Objekt šumu	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	Pozícia k.	0.076	80.026	5.547	9.494	2.643

Trojuholník	Pozícia k.	0.022	96.210	6.648	11.724	2.855
Priamky	Azimut k.	0.029	238.456	12.344	29.591	4.086
Trojuholník	Azimut k.	0.030	4255.725	31.205	230.286	4.425
Priamky	BB	0.020	188.712	4.772	17.252	1.301
Trojuholník	BB	0.068	116.442	4.110	9.598	1.552

Tabuľka 19 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre *dataset up*.

Objekt šumu	Metóda lokalizácie	Metóda vylepšenia	Chyba
Pozícia k.	Priamky	Primitívny centroid	0.737
Pozícia k.	Trojuholník	Primitívny centroid	0.360
Pozícia k.	Priamky	Zhlukovanie	1.174
Pozícia k.	Trojuholník	Zhlukovanie	0.391
Azimut k.	Priamky	Primitívny centroid	1.858
Azimut k.	Trojuholník	Primitívny centroid	13.695
Azimut k.	Priamky	Zhlukovanie	1.736
Azimut k.	Trojuholník	Zhlukovanie	0.796
BB	Priamky	Primitívny centroid	5.497
BB	Trojuholník	Primitívny centroid	0.164
BB	Priamky	Zhlukovanie	0.525
BB	Trojuholník	Zhlukovanie	1.230

Tabuľka 20 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre *dataset up*.

Dataset straight-left

Metóda	Objekt šumu	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	Pozícia k.	0.240	176.363	8.632	16.720	3.362
Trojuholník	Pozícia k.	0.087	231.347	13.062	23.313	5.160
Priamky	Azimut k.	0.124	5223.918	44.043	326.004	7.804
Trojuholník	Azimut k.	0.088	1788.535	25.469	109.205	7.513
Priamky	BB	0.232	266.698	7.627	22.367	2.541
Trojuholník	BB	0.167	394.652	11.705	37.454	3.154

Tabuľka 21 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre *dataset straight-left*.

Objekt šumu	Metóda lokalizácie	Metóda vylepšenia	Chyba
Pozícia k.	Priamky	Primitívny centroid	1.233
Pozícia k.	Trojuholník	Primitívny centroid	2.781
Pozícia k.	Priamky	Zhlukovanie	0.942
Pozícia k.	Trojuholník	Zhlukovanie	3.081
Azimut k.	Priamky	Primitívny centroid	5.521
Azimut k.	Trojuholník	Primitívny centroid	4.550
Azimut k.	Priamky	Zhlukovanie	0.216
Azimut k.	Trojuholník	Zhlukovanie	4.810
BB	Priamky	Primitívny centroid	3.570
BB	Trojuholník	Primitívny centroid	0.446
BB	Priamky	Zhlukovanie	0.998

BB	Trojuholník	Zhlukovanie	2.740
----	-------------	-------------	-------

Tabuľka 22 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre *dataset straight-left*.

Dataset straight-right

Metóda	Objekt šumu	Min.	Max.	Priemer	Smer. odch.	Medián
Priamky	Pozícia k.	0.044	278.208	18.418	33.266	7.905
Trojuholník	Pozícia k.	0.118	501.444	19.289	43.869	6.072
Priamky	Azimut k.	0.123	1365.904	27.945	81.410	9.490
Trojuholník	Azimut k.	0.087	2320.623	46.318	181.372	10.820
Priamky	BB	0.064	962.414	20.768	80.717	4.541
Trojuholník	BB	0.025	935.226	16.433	62.281	3.537

Tabuľka 23 Štatistika chybovosti (v metroch) jednotlivých odhadov s aplikovaným šumom pre *dataset straight-right*.

Objekt šumu	Metóda lokalizácie	Metóda vylepšenia	Chyba
Pozícia k.	Priamky	Primitívny centroid	1.562
Pozícia k.	Trojuholník	Primitívny centroid	0.678
Pozícia k.	Priamky	Zhlukovanie	0.398
Pozícia k.	Trojuholník	Zhlukovanie	0.024
Azimut k.	Priamky	Primitívny centroid	8.960
Azimut k.	Trojuholník	Primitívny centroid	4.087
Azimut k.	Priamky	Zhlukovanie	5.254
Azimut k.	Trojuholník	Zhlukovanie	9.858
BB	Priamky	Primitívny centroid	8.525
BB	Trojuholník	Primitívny centroid	3.874
BB	Priamky	Zhlukovanie	1.631
BB	Trojuholník	Zhlukovanie	0.333

Tabuľka 24 Štatistika chybovosti (v metroch) výsledného odhadu s aplikovaným šumom pre *dataset straight-right*.

Aplikácia šumu na niektorý z prvkov výpočtu ukázala očakávané dôsledky. Najlepší odhad sa relatívne nezmenil, v niektorých prípadoch bol ešte lepší. Dramaticky narástol odhad najhorší. Nakoľko bol použitý pseudonáhodný generátor čísel (angl. *pseudo-random number generator* (PSRNG)) pre normálne rozdelenie pravdepodobnosti, tak je šanca, že sme získali aj extrémnu modifikáciu hodnoty. Distribúcia týchto extrémnych bodov odpovedá vedomostiam o normálnom rozdelení pravdepodobnosti.

Dôležitým faktorom je uvedomiť si medián, ktorý je nie na extrémne hodnoty citlivý. Hodnota mediánu nebola dramaticky ovplyvnená (v porovnaní s maximálnou chybou a odchýlkou, napríklad), aj keď chyba narástla v rozsahu od 1 do 5 metrov.

3.4.4 Výsledky testovania pre reálny dataset

Dataset zilina – objekty s dôveryhodnou skutočnou pozíciou

Metóda	Min.	Max.	Priemer	Smer.odch.	Medián
Priamky	0.149	12.039	4.382	1.813	4.281
Trojuholník	0.689	6.998	3.746	1.687	3.340

Tabuľka 25 Štatistika chybovosti (v metroch) jednotlivých odhadov pre *dataset zilina* (časť 1).

Metóda lokalizácie	Metóda vylepšenia	Min.	Max.	Priemer	Smer.odch.	Medián
Priamky	Primitívny centroid	1.403	7.047	3.856	1.585	4.457
Trojuholník	Primitívny centroid	0.689	6.998	3.746	1.687	3.340
Priamky	Zhlukovanie	1.589	7.513	4.503	1.753	4.062
Trojuholník	Zhlukovanie	0.643	7.209	4.221	1.728	3.810

Tabuľka 26 Štatistika chybovosti (v metroch) výsledného odhadu pre *dataset zilina* (časť 1).

Dataset zilina – objekty s nízkou dôveryhodnosťou v skutočnú pozíciu

Metóda	Min.	Max.	Priemer	Smer.odch.	Medián
Priamky	0.035	7981.244	22.264	161.384	7.254
Trojuholník	0.105	10215.531	27.374	197.404	7.872

Tabuľka 27 Štatistika chybovosti (v metroch) jednotlivých odhadov pre *dataset zilina* (časť 2).

Metóda lokalizácie	Metóda vylepšenia	Min.	Max.	Priemer	Smer.odch.	Medián
Priamky	Primitívny centroid	3.004	44.565	13.651	11.744	9.241
Trojuholník	Primitívny centroid	3.584	56.774	16.561	15.660	9.152
Priamky	Zhlukovanie	0.824	17.870	5.110	2.442	5.032
Trojuholník	Zhlukovanie	1.664	17.829	5.961	2.316	6.682

Tabuľka 28 Štatistika chybovosti (v metroch) výsledného odhadu pre *dataset zilina* (časť 2).

Reálny *dataset* vykazoval vysoké rozpätie chybovosti v dôsledku nepresného GPS vzorkovania, o ktorom sme sa už zmienili (prípadne iných faktorov v rámci spracovania dát (interpolácia atď.)).

Ako najdôležitejšiu z hľadiska praktickej aplikovateľnosti považujeme tabuľku 26, ktorá sumarizuje výsledky najlepších odhadov pre reálny *dataset*, kde nie sú objekty, ktoré boli problematické pri zameriavaní polohy, a teda je vyššia šanca, že sa nám podarilo zachytiť správnu polohu.

Metóda je schopná určiť odhad polohy objektu pod 1 meter, v priemere je chyba okolo 4 metrov, spoločne s mediánom. Na RD s dôveryhodnými polohami dopravných

značení sme boli schopný získať chybovosť v najhoršom prípade okolo 7 metrov. Metóda na základe výsledkov vyzerá byť použiteľná pre praktické účely.

3.4.5 Testovanie chybovosti vzhľadom na vzdialenosť od objektu

Motiváciou pre tento typ testu bol fakt, že chyba odhadu súvisí so vzdialenosťou objektu od kamery. Ukážeme, ako sa mení minimálna, maximálna a priemerná chyba spoločne so smerodajnou odchýlkou a mediánom vzhľadom na vzdialenosť kamery od objektu v čase odhadovania pozície.

Úlohou je realizovať výpočty všetkých možných dvojíc snímok, na ktorých je objekt viditeľný, a zaznamenať chybu odhadu. Budeme testovať, ako sa menia hodnoty vyššie uvedených štatistických ukazovateľov pri zmene prahovej hodnoty pre minimálnu a maximálnu vzdialenosť objektu od kamery pre konkrétnu dvojicu snímok.

Konkrétne, nech $d(f_1)$ je vzdialenosť objektu od kamery na prvej snímke, a nech $d(f_2)$ platí analogicky pre druhú snímku. Pre prahové hodnoty t_{min} a t_{max} , ktoré budú predstavovať parametricky určené hranice pre minimálnu a maximálnu vzdialenosť, vykonáme štatistickú analýzu odhadov polohy len pre dvojice snímok, pre ktoré platí, že $d(f_1)$ a $d(f_2)$ sú z intervalu (t_{min}, t_{max}) . Takýmto postupom rozdelíme množinu dvojíc do niekoľkých disjunktných podmnožín, nad ktorými vykonáme štatistickú analýzu.

Idea je parametre t_{min} a t_{max} modifikovať iteratívne v rozsahoch hustejších pri vzdialenostiach nižších, a v redších pri vyšších vzdialenostiach, nakoľko pri vzdialenosti nad 100 metrov nie sú informácie tak relevantné ako pri vzdialenostiach od 10 do 20 metrov. Výsledky nebudeme uvádzať pre jednotlivé *datasety* ale budú spojené dohromady. Dôvodom je rozsah práce a aj samotná myšlienka, že získame všeobecný prehľad o presnosti metód naprieč niekoľkými typmi ciest, čo je v realite bežné.

Pri výpočtoch boli zmiernené filtrovania aby sme získali aj rizikové kombinácie vzhľadom na vysoké vzdialenosti, ktorým sme sa obyčajne vyhli. Toto je dôvodom, prečo je maximálna chyba vyššia ako maximálna chyba, ktorú je možné nájsť pri jednotlivých SD. Naopak, nakoľko sme výhradne používali len kombinácie snímok, ktoré spĺňajú vyššie stanovenú podmienku o vzdialenosti objektu od kamery, tak sme aj zamietli možnosť získať odhady najlepšie, preto sa nám nepodarilo získať rovnakú minimálnu chybu, nakoľko tie boli často dôsledkom situácie, kedy rozdiel vo veľkostiach BB ako aj uhloch (ktoré súvisia so vzdialenosťou objektu) bol pre naše metódy ideálny, ale v takomto prípade by vzdialenosti objektu od kamery presahovali hranice stanoveného intervalu.

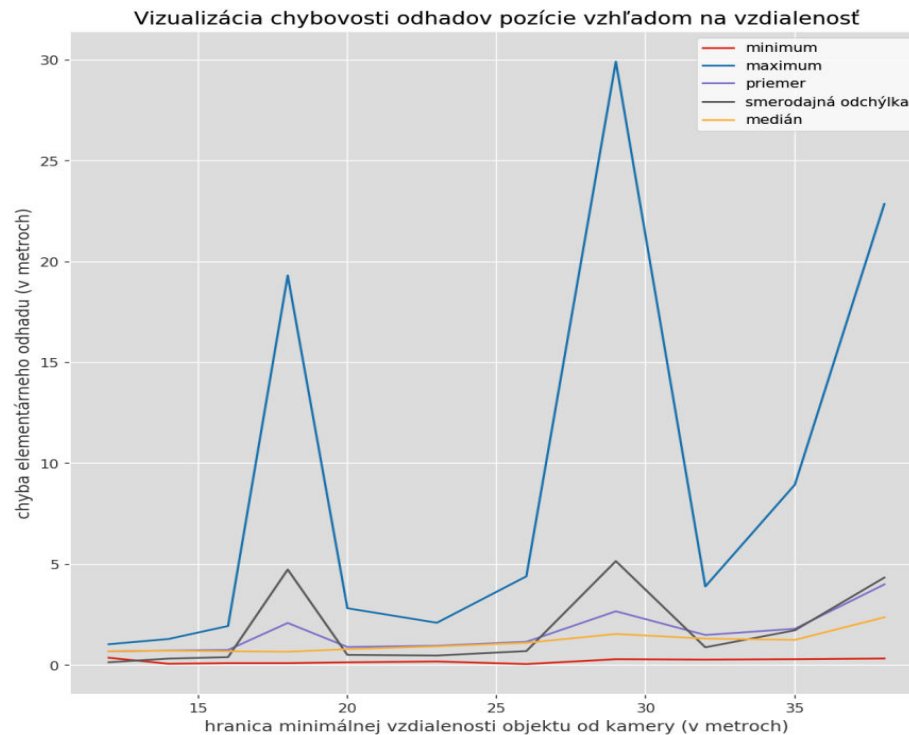
Tento test predstavuje len doplnok aby sme získali prehľad o vplyve vzdialenosti. Je možné z neho usúdiť, ako sa metódy správajú v najhoršom prípade ak máme k dispozícii len pohľad zo vzdialenosti nad niekoľko metrov. Fakt, že sa jedná o najhorší prípad je zrejmý z toho, že v praxi by sme pracovali so všetkými kombináciami snímok, a získali buď rovnaký alebo lepší výsledok.

V tomto prípade sme určili relatívne nerealistické podmienky použitia len za účelom získania jasnej predstavy o vplyve vzdialenosti objektu od kamery. Takáto forma testovania nám pripadala najjednoduchšia bez ujmy na platnosti našich záverov.

Interval vzdialenosti	Minimum	Maximum	Priemer	Smer. odch.	Medián	Počet vzoriek
(12, 14)	0.374	1.045	0.689	0.150	0.694	24
(14, 16)	0.076	1.303	0.731	0.332	0.722	46
(16, 18)	0.108	1.947	0.761	0.405	0.696	62
(18, 20)	0.108	19.321	2.098	4.747	0.672	84
(20, 23)	0.147	2.828	0.908	0.517	0.799	186
(23, 26)	0.189	2.108	0.969	0.485	0.936	174
(26, 29)	0.062	4.408	1.164	0.703	1.114	198
(29, 32)	0.303	29.917	2.678	5.163	1.552	120
(32, 35)	0.282	3.907	1.503	0.891	1.324	102
(35, 38)	0.303	8.950	1.809	1.733	1.259	102
(38, 40)	0.337	22.863	4.008	4.347	2.380	102
(40, 45)	0.283	42.225	7.622	6.076	6.522	310
(45, 50)	0.185	692.129	76.801	112.901	44.203	270
(50, 55)	10.001	481.204	62.308	52.409	52.258	290
(55, 60)	0.482	313.163	26.381	25.444	25.102	270
(60, 70)	0.303	54.690	10.193	7.105	9.399	1180
(70, 80)	0.887	51.286	20.972	6.719	19.351	602
(100, 110)	7.531	44.244	31.431	5.160	31.771	242
(110, 120)	1.121	840.453	20.167	45.574	13.986	882
(120, 130)	17.762	44041.983	1363.931	4140.984	402.401	840
(130, 140)	0.195	4898.681	107.471	264.718	84.663	840
(140, 150)	1.260	5668.258	219.013	809.118	40.951	574
(150, 160)	0.513	164.456	51.998	38.888	42.149	488

Tabuľka 29 Štatistika chybovosti odhadov pozície vzhľadom na vzdialenosť objektu od kamery.

V grafe nižšie uvádzame len údaje do vzdialenosti 40 metrov, nakoľko maximálna chyba prudko narastala pri vyšších vzdialenostiach a prispôsobenie škálovania grafu rezultovalo v problematickú interpretáciu hodnôt.



3-12 Vizualizácia chybovosti odhadov pozície vzhľadom na vzdialenosť objektu od kamery.

Je evidentné, že minimálnu chybu odhadu, ktorú môžeme považovať za výbornú je možné získať aj na vzdialenosť 40 metrov. Vysokú variabilitu má chyba maximálna, ktorá priamo ovplyvňuje priemer, a ten rovnako vykazuje trend „skokovej chybovosti“ (spoločne so smerodajnou odchýlkou).

Tento trend nastáva v dôsledku prítomnosti *outliers*. Pri takomto spôsobe testovania, aký sme zvolili, je ich výskyt očakávaný. Nakoľko uhly, pod ktorými je objekt viditeľný a veľkosti BB sú podobné, je zrejmé, že je možné získať odhad veľmi nepresný len vplyvom týchto faktorov a náhody.

Medián elementárneho odhadu postupne rastie so zvyšujúcou sa hranicou vzdialenosti. Ako je z grafu evidentné, použitie metódy do 30 metrov je vhodné, ale ideálne je, aby sa kamera dostala do bezprostrednej blízkosti objektu, napríklad, aby vozidlo prešlo okolo dopravného značenia. Tento predpoklad nie je náročné splniť, nakoľko pri snahe odhadovať pozície objektu je v záujme používateľa aby získal vhodné zábery objektu záujmu.

Záver

V práci boli navrhnuté, implementované a experimentálne testované 2 metódy pre odhad pozície statického objektu vzhľadom na pozíciu pohybujúcej sa kamery. Metódy boli použité pre odhadovanie pozície anotovaného objektu viditeľného vo video sekvencii, ktorá bola zaznamenaná jednou kamerou. Pre fungovanie týchto metód je postačujúca len jedna kamera, čo je najväčšou výhodou použitia. Jedinou potrebnou informáciou o kamere je horizontálne uhlové zorné pole, ktoré nie je náročné zistiť.

Na základe predpokladov je výsledné riešenie vhodné pre použitie v prípadoch, kedy robustná stereo vícia pomocou špecializovaných kamier nie je možná. Prístup kladie relatívne malé požiadavky na hardvér, vstupné dáta a spôsob ich získania. Za istých okolností dokáže získať uspokojivý výsledok a riešenie pomocou oboch metód odhadu pozície objektu je tak vhodné pre použitie v praxi. Predpokladom je, aby sa kamera dostala do vzdialenosti od objektu pod 30 metrov pre použiteľné výsledky. Ideálne je, ak kamera okolo objektu prejde.

Metódy fungujú na báze vykonávania elementárnych odhadov pozície objektu záujmu so všetkých dostupných snímok, na ktorých je objekt viditeľný. Na relatívne vysoký počet odhadov neskôr aplikujeme rovnako dve metódy, pomocou ktorých určíme výsledný odhad pozície objektu. Jedna metóda predstavuje triviálnu kalkuláciu centroidu všetkých bodov, pričom druhá metóda funguje na princípe odstraňovania potenciálne chybných odhadov, ktoré môžu interferovať s klasickým výpočtom centroidu, ktorý je veľmi citlivý na extrémne hodnoty.

Umelé vytvorené syntetické dáta ukázali presnosť, akú je možné dosiahnuť v laboratórnych podmienkach. Elementárne odhady sa pohybovali v chybovosti od 4 mm až po niekoľko metrov. Dôležitým faktorom je ale výsledný odhad, ktorý bol určený na základe odhadov elementárnych. Všeobecné tvrdenie, ktoré platí pre syntetické dáta, pre obidve metódy výpočtu elementárneho a aj finálneho odhadu pozície znie: minimálna chybovosť je 30 cm, priemerná 70 cm a medián 60 cm. V skratke, je možné určiť polohu objektu do jedného metra na rôznych typoch náročnosti terénu, v ktorom sa objekt nachádzal ako aj pohybu kamery, ktorou bol snímaný.

Rozdiely v aplikovaných metódach ako aj ďaleko vyššie chybovosti boli zaznamenané pri reálnych dátach, ktoré nám ukázali potenciálne nástrahy pri aplikácii

našich metód. Prvým z nich je nepresné vzorkovanie GPS. Tu nastali aj rozdiely aj v jednotlivých metódach odhadu pozície, ale badateľné len v prípade najlepšieho odhadu. Priemer a medián boli opätovne veľmi blízke. Minimálna chyba odhadu je 60cm / 140cm pre obidve metódy, pričom priemerná chyba a medián sa pohybujú okolo 4 metrov. Rovnako aj rozdielne metódy pre určenie výsledného odhadu z elementárnych vykazujú malú, ale stále badateľnú odchýlku vo výslednom odhade.

Metódy je možné použiť ak existuje už zaznamenaná video sekvencia bez možnosti opätovného nahrávania, v ktorej je potrebné odhadnúť pozíciu, prípadne aspoň vzdialenosť k nejakým objektom. Naše metódy preukázali potenciál pre prínosnú, implementačne nenáročnú praktickú aplikáciu len za použitia jednej kamery, ktorá je v dnešnej dobe prítomná všade, či už v rámci mobilného zariadenia, tabletu atď.

Na základe teoretickej chybovosti do jedného metra môžeme usúdiť, že po matematickej stránke je aj napriek relatívne triviálnej formulácii tento postup účinný. S chybovosťou v priemere okolo 4 metrov v rámci praktickej aplikácie môžeme tvrdiť o možnosti získať nie veľmi presnú, avšak uspokojivú a rozhodne využiteľnú informáciu o pozícií skúmaného objektu s minimálnym množstvom vstupných informácií.

Predpokladáme, že teoreticko-praktická téma tejto práce prinesie do budúcnosti úžitok, či už v zmysle dodatočného skúmania vylepšovania odhadov samotných metód, alebo rozšírenia týchto metód do 3 rozmerov v snahe pokúsiť sa odhadovať aj vertikálnu pozíciu (Z súradnicu) konkrétného objektu.

Vznikla tu aj otázka kvantifikovania dôveryhodnosti v odhad pozície, ktorý by mohol v sebe niesť viacero faktorov, akými je rozlíšenie obrázka, uhol, pod ktorými sú objekty viditeľné a pod. Dokonca je možnosť pri získaní pôvodných, „surových“ dát z GPS zariadenia extrahovať aj počet dostupných satelitov, na základe ktorých bola pozícia určená. V takomto prípade by mohli byť odhady vážené takto vytvoreným koeficientom, a teda silne chybové, extrémne hodnoty by získavali relatívne malú váhu v ideálnom stave, ak by sme dokázali korektne identifikovať a matematicky vyjadriť možné negatívne vplyvy.

Táto téma nie je v tomto stave uzavretá, avšak môžeme tvrdiť, že je v stave praktickej použiteľnosti, ktorú sme patrične demonštrovali. Dúfame, že budúci výskum v tejto oblasti prinesie nápady pre možné zlepšenie našich postupov v rámci problému odhadovania pozície statického objektu z pohyblivej kamery.

Zoznam použitej literatúry

- [1] *Holzmann Clemens, Hochgatterer Matthias*: **Measuring Distance with Mobile Phones Using Single-Camera Stereo Vision**. University of Applied Sciences Upper Austria.
- [2] *Parker Andrew, Smith Jackson, Krug Kristine*: **Neural architectures for stereo vision**.
- [3] *Tsai Victor, Chang Chun-Ting*: **Three-dimensional positioning from Google street view panoramas**. Department of Civil Engineering, National Chung Hsing University, 2012.
- [4] *Bénallal Mohamed, Meunier Jean*: **A simple algorithm for object location from a single image without camera calibration**.
- [5] *Chen Chi-Farn, Chen Min-Hsin*: **Object tracking and positioning on video images**. Center for Space and Remote Sensing Research, National Central University.
- [6] *Prematunga H., Dharmaratne Anuja*: **Finding 3D Positions from 2D Images Feasibility Analysis**. University of Colombo School of Computing.
- [7] *Saxena Ashutosh, Chung Sung, Ng Andrew*: **Learning Depth from Single Monocular Images**. Computer Science Department, Stanford University.
- [8] *Kim Insu, Yow Kin*: **Object Location Estimation from a Single Flying Camera**. GIST College, Gwangju Institute of Science and Technology.
- [9] „**Random Dot Stereogram**“ (*on-line*):
https://en.wikipedia.org/wiki/Random_dot_stereogram
- [10] „**Depth Map**“ (*on-line*):
https://en.wikipedia.org/wiki/Depth_map
- [11] „**Active Pixel Sensor**“ (*on-line*):
https://en.wikipedia.org/wiki/Active_pixel_sensor
- [12] „**Pinhole Camera Model**“ (*on-line*):
https://en.wikipedia.org/wiki/Pinhole_camera_model
- [13] „**Azimut** (astronómia)“ (*on-line*):

[https://sk.wikipedia.org/wiki/Azimut_\(astron%C3%B3mia\)](https://sk.wikipedia.org/wiki/Azimut_(astron%C3%B3mia))

[14] „Azimuth“ (*on-line*):

<https://en.wikipedia.org/wiki/Azimuth>

[15] „Zorné pole“ (*on-line*):

[https://sk.wikipedia.org/wiki/Zorn%C3%A9_pole_\(optika\)](https://sk.wikipedia.org/wiki/Zorn%C3%A9_pole_(optika))

[16] „Blender“ (*on-line*):

<https://www.blender.org/>

[17] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani: **An Introduction to Statistical Learning (with Applications in R)** | ISBN: 978-1-4614-7137-0

[18] “Field of view” (*on-line*):

https://en.wikipedia.org/wiki/Field_of_view

[19] Needham Joseph: **Science and Civilisation in China**. Cambridge University Press, 1962 vol. 4, pt. 1

[20] Gil-Jiménez Pedro, Gómez-Moreno Hilario, López-Sastre Roberto and Maldonado-Bascón Saturnino: **Geometric bounding box interpolation: an alternative for efficient video annotation**. Gil-Jiménez et al. EURASIP Journal on Image and Video Processing (2016) 2016:8.

[21] „Funkcia **atan2**(y, x)“ (*on-line*):

<https://en.wikipedia.org/wiki/Atan2>

Prílohy

Príloha A: Obsah DVD

Priložené DVD obsahuje:

- Práca v elektronickej podobe (formát PDF).
- Zdrojové kódy testovacej aplikácie (programovací jazyk C#, projekt pre aplikáciu Microsoft Visual Studio).
- Zdrojové kódy pomocných nástrojov pre transformácie dátových množín (programovací jazyk Python).
- Zdrojové kódy pomocných nástrojov pre štatistické analýzy výsledkov (programovací jazyk Python, súbory vo formáte Jupyter Notebook).
- CSV (Comma Separated Values) súbory obsahujúce výsledky, na základe ktorých boli vytvorené štatistické analýzy a závery.
- Všetky vytvorené syntetické dátové množiny.
- Všetky vytvorené reálne dátové množiny.
- Podrobný popis obsahu DVD ako aj ukážka použitia aplikácie pre získavanie výsledkov (formát PDF).