

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

VIZUÁLNY SYSTÉM PRE INTERAKCIU
ĽUDSKÉHO UČITEĽA S ROBOTOM
DIPLOMOVÁ PRÁCA

2021
ANGELIKA FEDÁKOVÁ

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

VIZUÁLNY SYSTÉM PRE INTERAKCIU
ĽUDSKÉHO UČITEĽA S ROBOTOM
DIPLOMOVÁ PRÁCA

Študijný program: Aplikovaná informatika
Študijný odbor: Aplikovaná informatika
Školiace pracovisko: Katedra aplikovanej informatiky
Školiteľ: Ing. Viktor Kocur

Bratislava, 2021
Angelika Fedáková



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Angelika Fedáková
Študijný program: aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Vizualný systém pre interakciu ľudského učiteľa s robotom
Visual system for interaction of a human teacher with a robot

Anotácia: Toto zadanie je súčasťou projektu interakcie ľudského učiteľa s robotom. Robot pri tejto interakcii manipuluje jednoduchými objektmi na základe pokynov od ľudského učiteľa. Pre tento účel je tak vhodné aby robot dokázal správne detegovať pozíciu jednoduchých, objektov pomocou, učiteľovej ruky a svojho robotického ramena pomocou svojej RGB-D kamery.

Cieľ: Cieľom tejto práce je navrhnúť, implementovať a otestovať systém ktorý na základe vstupných RGB-D dat z Intel RealSense kamery deteguje pozíciu jednoduchých objektov, učiteľovej ruky a robotického ramena. Súčasťou práce bude prehľad existujúcich riešení detekcie objektov v RGB-D snímkach. Navrhnutý algoritmus bude vyhodnotený v kontexte prebiehajúceho projektu interakcie ľudského učiteľa s robotom.

Vedúci: Ing. Viktor Kocur
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: prof. Ing. Igor Farkaš, Dr.
Dátum zadania: 07.10.2020

Dátum schválenia: 08.10.2020

prof. RNDr. Roman Ďurikovič, PhD.
garant študijného programu

.....
študent

.....
vedúci práce

Pod'akovanie: Ďakujem všetkým, ktorí ma pri tvorbe tejto práce podporovali.

Abstrakt

Abstrakt práce.

Klíčové slová: klucove slova

Abstract

English abstract.

Keywords: keywords

Obsah

Úvod	1
1 Technické pozadie práce	7
Záver	9
Prílohy	13

Zoznam obrázkov

Zoznam tabuliek

Úvod

S automatickými alebo poloaautomatickými robotickými prístrojmi sa v dnešnej dobe stretávame čoraz častejšie. Na to, aby na reálny svet roboty dokázali reagovať a komunikovať s ním je často potrebné aby dokázali rozpoznávať čo sa okolo nich deje prostredníctvom kamier. V tejto práci sme sa venovali videniu u robota NICO a rozpoznávaniu detských drevených kociek v jeho zornom poli.

Oblasť počítačového videnia je veľmi skúmaná oblasť. Je to najmä vďaka jej širokému využitiu v riešení problémov reálneho sveta. Jednoduché detektory môžeme nájsť kdekoľvek, od priemyselných kamier až po mobilné telefóny. Oblasť objektovej detekcie je jednou z úloh počítačového videnia. Je to proces, kedy sa v obraze snažíme nájsť a označiť jednotlivé požadované objekty. Na túto úlohu môžeme použiť základné metódy počítačového videnia, ako sú napríklad hranové detektory, alebo princípy hlbokého učenia a takzvané konvolučné neurónové siete. Tieto siete majú ale jeden základný problém. Na ich použitie je potrebné tréningovanie a na to je nutnosť mať k dispozícii čo najväčšie množstvo ošetrovaných dát. Ručný zber dát a ich anotácia je takmer vždy časovo náročná a pri väčších datasetoch a obmedzených zdrojoch môže byť nespĺniteľná. Z toho dôvodu sme sa v tejto diplomovej práci venovali testovaniu procesu, v ktorom sme na tréningovanie CNN pre detekciu drevených objektov využili iba počítačovo vygenerované obrazy.

Objektová detekcia

Počítačové videnie

Pre človeka je vnímanie 3D priestoru reálneho sveta a objektov v ňom prirodzené a jednoduché. Ľudia dokážu od seba rozlišovať jednotlivé objekty a k tomu aj ich farby, textúry, či veľkosti pomocou malých náznakov svetla, či tieňov. Pochopenie ľudského vnímania sveta je veľmi zložitá a o jeho pochopenie, respektívne následné namodelovanie v digitálnom prostredí sa snaží odvetvie s názvom počítačové videnie.

Ultimátnym cieľom počítačového videnia je z digitálneho obrazu získať plné sémantické pochopenie zobrazovanej reality. Aj keď sme sa v mnohých oblastiach, ako je napríklad detekcia ľudí, či tvárí v posledných rokoch výrazne posunuli, plné sémantické pochopenie zostáva stále v nedohľadne. Problémom je najmä komplexita informácií, ktorú človek vníma pri sledovaní obrazu, ktoré z obrazu často nie sú jasné a je nutné poznať isté súvislosti. [2]

Objektová detekcia

Problém hľadania inštancií objektov istých tried v digitálnom obraze nazývame objektová detekcia. Jeho riešenie nám poskytuje základné informácie o rozložení objektov v obraze. Poznáme dva typy úloh objektovej detekcie. Prvou je všeobecná detekcia objektov, pri ktorej sa snažíme rozpoznať objekty rôznych tried jedným procesom. Druhý typ úlohy je aplikačná detekcia objektov, kde je hľadanie zúžené na konkrétne triedy objektov, ako napríklad ľudské tváre, či prechody pre chodcov. [9]

Poznáme viacero prístupov riešenia tohto problému, ktoré môžeme zaradiť do dvoch hlavných skupín. Prvou skupinou sú tradičné metódy počítačového videnia, ktoré sú založené na ručnom určení príznakov objektov a ich následnom hľadaní v obraze. Druhá skupina metód sú založené na hlbokom učení, konkrétne konvolučných neurónových sieťach, ktoré si príznaky jednotlivých objektov určujú samy.

Tradičné metódy počítačového videnia

Väčšina týchto prístupov je založených na ručne určených príznakoch. Ich výhodou je, že pri trénovaní nepotrebujeme veľké množstvo dát. Určenie vhodných príznakov však môže byť veľmi náročné. Zatiaľčo pri jednoduchých geometrických objektoch sú tieto metódy pomerne presné a veľmi rýchle, pri zložitejších objektoch, ako sú napríklad mačky a psy už určiť vhodné príznaky nie je také jednoduché. Pozrime sa na pár základných algoritmov, ktoré výrazne ovplyvnili históriu objektovej detekcie.

1. Viola Jones detektor - prvý detektor ľudských tvárí v reálnom čase, ktorý v svojej dobe rýchlosťou výrazne predbehol ostatné algoritmy. Svoju rýchlosť dosiahol najmä vďaka využitiu integrálnych obrazov, algoritmický výber Haárových filtrov a viackroková detekcia.
2. Histogram orientovaných gradientov - je príznakový deskriptor. V svojom období prispel najmä k riešeniu problému škálovej a tvarovej invariantnosti pri objektovej detekcii. Je založený na rozdelení obrazu do uniformnej mriežky. Každéj bunke tejto mriežky sú následne priradené počty nájdených gradientov v rôznych smerov.

Metódy založené na hlbokom učení

Tieto metódy priniesli do problému objektovej detekcie ale aj iných problémov počítačového videnia veľký pokrok. Sú zväčša založené na konvolučných neurónových sieťach (CNN). Príznaky si siete určujú samy v procese učenia. Na to je potrebné veľké množstvo oánotovaných dát, pomocou ktorých sa sieť postupne učí, ako objekty rozpoznať. Tieto siete môžeme taktiež rozdeliť na dve skupiny podľa ich prístupu k hľadaniu objektov.

1. dvojkrokové algoritmy - v prvom kroku rýchlejšie určujú približné širšie oblasti a v druhom kroku ich spresňujú. Jedným zo zástupcov týchto prístupov je RCNN, ktorej architektúra je pomerne jednoduchá. V prvom kroku extrahuje potenciálne oblasti s objektami a v druhom kroku sú tieto oblasti preškálované do správnej veľkosti a sú poslané ako vstup do CNN, ktorá následne extrahuje príznaky. Nakoniec je na týchto príznakoch využitá metóda SVM, ktorá určí, či sa v danej oblasti nachádza objekt a ak áno, o objekt akej triedy ide. Nevýhodou RCNN je ale pomalé opätovné určovanie príznakov, ktoré bolo adresované inými dvojkrokovými algoritmami.
2. jednokrokové algoritmy - finálne objekty hľadajú v jednom kroku. Jedným zo zástupcov tohto prístupu je YOLO, ktorý bol prvým jednokrokovým riešením

v dobe hlbokého učenia. Do CNN je vstupom narozdiel od dvojkrokových algoritmov posielaný celý obraz. CNN obraz rozdeľuje na oblasti a určuje pravdepodobnosti objektov jednotlivých oblastí naraz. Aj keď je oproti dvojkrokovým algoritmom rýchlejší, zaostáva v presnosti lokalizácie a to najmä pri menších objektoch. [9]

Kapitola 1

Technické pozadie práce

Robot NICO

Meno tohto robota pochádza z anglického názvu Neuro-Inspired COmpanion, čo by sme mohli voľne preložiť ako neuro-inšpirovaný spoločník, navrhnutý na Universität of Hamburg. Je to robotický humanoid určený pre vývojové účely. Cieľom je adresovať vnemové a interakčné schopnosti robota potrebné pri výskume v neuro-kognitívnej oblasti.

Schopnosti

NICO obsahuje LED maticové displeje v oblasti obočia a úst, čo mu umožňuje vyobrazovať rôzne emócie. Obsahuje 2 kamery so 4K rozlíšením a schopnosťou nahrávania v 60fps v oblasti očí. Dve mikrofóny nachádzajúce sa v oblasti uší umožňujú využívať stereoaudio. RH4D manipulátory v oblasti rúk umožňujú pokročilé možnosti uchopovania a prijímania zložitých senzorických vnemov z prostredia.

Záver

Záver

Literatúra

- [1] Krystof Litomisky. Consumer rgb-d cameras and their applications. *Rapport technique, University of California*, 20:28, 2012.
- [2] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- [3] Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. In *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pages 23–30. IEEE, 2017.
- [4] Jonathan Tremblay, Thang To, Balakumar Sundaralingam, Yu Xiang, Dieter Fox, and Stan Birchfield. Deep object pose estimation for semantic robotic grasping of household objects. *arXiv preprint arXiv:1809.10790*, 2018.
- [5] Yilin Wang and Jiayi Ye. An overview of 3d object detection. *arXiv preprint arXiv:2010.15614*, 2020.
- [6] Isaac Ronald Ward, Hamid Laga, and Mohammed Bennamoun. Rgb-d image-based object detection: From traditional methods to deep learning techniques. In *RGB-D Image Analysis and Processing*, pages 169–201. Springer, 2019.
- [7] Bin Yang, Wenjie Luo, and Raquel Urtasun. Pixor: Real-time 3d object detection from point clouds. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 7652–7660, 2018.
- [8] Xingyi Zhou, Dequan Wang, and Philipp Krähenbühl. Objects as points. *arXiv preprint arXiv:1904.07850*, 2019.
- [9] Zhengxia Zou, Zhenwei Shi, Yuhong Guo, and Jieping Ye. Object detection in 20 years: A survey. *arXiv preprint arXiv:1905.05055*, 2019.

Prílohy

Prílohy prace