

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

PRECHOD ZO SIMULÁCIE DO FYZICKEJ
REALITY V ROBOTIKE: PROBLÉMY A VÝZVY
BAKALÁRSKA PRÁCA

2024

JURAJ GAVURA

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

PRECHOD ZO SIMULÁCIE DO FYZICKEJ
REALITY V ROBOTIKE: PROBLÉMY A VÝZVY
BAKALÁRSKA PRÁCA

Študijný program: Aplikovaná informatika
Študijný odbor: Informatika
Školiace pracovisko: Katedra aplikovanej informatiky
Školiteľ: prof. Ing. Igor Farkaš, Dr.

Bratislava, 2024
Juraj Gavura



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:

Študijný program:

Študijný odbor:

Typ záverečnej práce:

Jazyk záverečnej práce:

Sekundárny jazyk:

Názov:

Anotácia:

Vedúci:

Katedra:

Vedúci katedry:

Dátum zadania:

Dátum schválenia:

garant študijného programu

.....
študent

.....
vedúci práce

Obsah

1	Úvod	1
2	Kvantifikácia rozdielu v realite	3
2.1	Identifikácia	3
2.2	Kvantifikácia	3
2.2.1	Testy	3
2.2.2	Zhrnutie	4
3	Riešenie problému rozdielu v realite	5
3.1	Známe metódy	5
3.1.1	Randomizácia domény	5
3.1.2	Adverzariálne učenie posilňovaním	5
3.1.3	Optimalizácia simulovaných prostredí	6
3.1.4	Transferové učenie	6
3.1.5	Použitie viacerých simulátorov súčasne	6
3.2	Nové metódy	7
3.2.1	Ladenie simulátoru	7
3.2.2	Diferencovateľná fyzika	7

Kapitola 1

Úvod

S každým novým rokom rastie dopyt po rôznych druhoch robotov. Či už ide o robotické rameno pracujúce s objektami, robot, ktorý sa potrebuje pohybovať v ťažkom teréne, alebo robot, ktorý využíva vizuál z kamery a musí rýchlo rozoznať, na čo sa pozerá. Predtým ako každý robot začne vykonávať svoje zadanie, musí prejsť určitou učiacou fázou. Jednou zo základných techník na cvičenie robota je učenie posilňovaním (angl. reinforcement learning). Toto učenie umožňuje vytvárať ovládače, ktoré sa dokážu naučiť optimálne správanie robota interakciou s okolím. Toto správanie, ktoré môžeme považovať ako stratégie (angl. policy), sú učené procesom pokus a omyl [5]. Z tohto dôvodu je tréning robota priamo vo fyzickej realite náročné časovo aj finančne, keďže je veľmi jednoduché robota počas tréningu poškodiť [4]. Preto je potrebné nájsť lepší spôsob na prácu s robotom vo vývoji.

Simulácia je široko využívaná vo svete robotiky vďaka tomu, že umožňuje roboty trénovať lacno, bez potreby prístupu k hardvéru a efektívne, keďže dokáže jednoducho generovať množstvo syntetických dát, ktoré podstatne zrýchlia proces učenia. Medzi ďalšie výhody simulácie patrí možnosť vykonávania viacerých inštancií simulácie paralelne a to všetko bez nutnosti ľudskej kontroly. Bohužiaľ, simulovanie učenia vo virtuálnej realite má aj svoje nevýhody. Z dôvodu nutnej aproximácie a abstrakcie niektorých fyzických javov, vznikajú nezrovnalosti medzi virtuálnou realitou v simulácii a fyzickou realitou. Simulácie tiež zvyknú vytvárať situácie, ktoré nemôžu nastať v našej realite. Tieto vlastnosti simulácie môžu značne komplikovať prestup zo simulácie do fyzickej reality. Tento problém vo svete robotiky nazývame rozdiel v realite (angl. reality gap) [3].

V tejto práci sa pozrieme na prieskum rôznych simulátorov a fyzikálnych motorov, kvantifikáciu rozdielu v realite, najviac využívané metódy prekročenia rozdielu v realite a na možné lepšie riešenia tohto problému.

Kapitola 2

Kvantifikácia rozdielu v realite

2.1 Identifikácia

Skôr ako sa pozrieme na to, ako by sme mohli kvantifikovať rozdiel v realite, musíme zistiť ako identifikujeme, že je rozdiel v realite vôbec prítomný. Predstavíme si priebeh učenia robota pomocou simulácie. Najprv sa v simulácii trénuje ovládač jednou z metód učenia a následne sa vo virtuálnej realite aj otestuje. Ak robot cieľ nesplní, ovládač sa trénuje ďalej. Ak naopak robot cieľ splní, test sa zopakuje na reálnom robotovi vo fyzickej realite. Ak nájdeme medzi výkonom robota v simulácii a výkonom robota vo fyzickej realite nejaké odlišnosti, nastal problém s prestupom do fyzickej reality a teda ide o rozdiel v realite [5]. Teraz keď sme identifikovali rozdiel v realite, môžeme sa ho pokúsiť nejakým spôsobom kvantifikovať.

2.2 Kvantifikácia

Ako uvádzajú Collins a kol. [3] vo svojej práci, táto kvantifikácia je možná do určitej miery. Otestovali najznámejšie simulátory s rôznymi fyzikálnymi motormi na sade troch testov pre prácu s robotickým ramenom Kinova Mico2 6DOF a porovnali pohyb v simulácii s pohybom v reálnom svete vyrátaním chyby pomocou euklidovskej vzdialenosti v závislosti od času. Ako simulátory si vybrali V-Rep, MuJoCo a PyBullet. Fyzikálne motory, ktoré tieto simulátory používajú sú Bullet, ODE, Vortex, Newton a MuJoCo. Testy obsahujú pohyb jednotlivých kĺbov ramena a interakciu s kockou.

2.2.1 Testy

Prvý test zahŕňal jednoduchú rotáciu jedného kĺbu ramena o 100 stupňov za 6 sekúnd. Najlepšie si vedli simulátory PyBullet a V-Rep s fyzikálnymi motormi Vortex a Newton napriek tomu, že v týchto simulátoroch rameno neoscillovalo tak, ako vo fyzickej realite.

Fyzikálny motor Bullet v simulátore V-Rep osciloval rovnako ako rameno vo fyzickej realite, ale prišiel do finálnej pozície o sekundu skôr.

V druhom teste sa striedavo rotovali dva kĺby ramena a hral v ňom veľkú úlohu pohyb proti gravitácii. Fyzickú realitu najlepšie kopíroval V-Rep s motormi Newton a Vortex a tesne za nimi PyBullet. Ostatné motory výrazne zaostávali a žiadnemu sa nepodarilo imitovať pohyb reálneho robota presne.

Tretí, finálny a najkomplexnejší test obsahoval interakciu s kockou. Táto interakcia spočívala v nastavení ramena do polohy vedľa kocky a následné jemné potlačenie kocky, teda tento finálny test zahŕňal pohyb viacerých kĺbov ramena. Žiadnemu simulátoru sa nepodarilo kopírovať pohyb kocky dokonalo. Najbližšie bol PyBullet, no iba z dôvodu, že kontakt s kockou nezaregistroval vôbec, rameno vošlo do kocky. MuJoCo a V-Rep s motorom Bullet boli najlepšie čo sa týka pohybu kocky, no zarotovali kocku o 90 stupňov, pričom v reálnom svete sa kocka zarotovala minimálne.

2.2.2 Zhrnutie

Z testov môžeme vidieť, že existujú značné rozdiely medzi jednotlivými simulátormi a fyzikálnymi motormi. Niektoré vynikali v jednom type scenáru, iné zase v druhom. Teda ideálne by bolo, keby bolo možné využiť jeden motor na jednu časť pohybu a druhý na inú. No očividné je, že simulácia interakcie dvoch pevných objektov potrebuje výrazné zlepšenie.

Kapitola 3

Riešenie problému rozdielu v realite

3.1 Známe metódy

Problém rozdielu v realite vo svete robotiky existuje už desiatky rokov a preto poznáme nemalý počet pokusov o riešenie tohto problému. Salvato a kol. [5] a tiež Collins a kol. [3] vo svojich prácach uvádzajú niekoľko potenciálnych riešení.

3.1.1 Randomizácia domény

Jedna z najpoužívanejších metód je randomizácia domény (angl. domain randomisation), ktorá už v minulosti mala pomerne dobré výsledky. V tomto prípade je trénovaný model vystavený náhodným vstupom a tým získa ovládač väčšiu robustnosť. Teda cieľom je, aby robot dokázal splniť svoju úlohu napriek rôznym prekážkam. Môžeme si to predstaviť na trénovaní robotického videnia, kde by model prijímal vstup ako náhodné hodnoty napríklad farby, tieňovania alebo pozície kamery, no stále by vedel rozoznať hľadaný objekt [3]. Mnoho ďalších metód na riešenie rozdielu v realite využíva náhodný výber. Adaptačná sieť s náhodným výberom (angl. randomised-to-canonical adaptation network) je spôsob randomizácie, pri ktorom sa pozorovania zozbierané vo fyzickej doméne spolu s pozorovaniami zo simulovanej domény zobrazia do spoločnej domény, ktorú nazývame kanonická doména [5]. Na podobnom princípe využívania dát zo simulácie spolu s dátami z reálneho sveta fungujú metódy ako adaptácia domény (angl. domain adaptation) a generatívne adverzariálne siete (angl. generative adversarial networks) [3].

3.1.2 Adverzariálne učenie posilňovaním

Inou metódou riešiacou problém rozdielu v realite je adverzariálne učenie posilňovaním (angl. adversarial reinforcement learning). Pri tomto spôsobe učenia sa trénujú dva ovládače, jeden sa nazýva protagonista a druhý antagonista. Úlohou druhého ovládača

je generovať stále zložitejšie a zložitejšie modely na zvládnutie cieľu robota. Na túto metódu nadväzuje algoritmus nazývaný robustné adverzariálne učenie posilňovaním. V tomto prípade je dodatočná spoľahlivosť zaručená tým, že antagonista navyše vnáša do modelu prostredia destabilizujúce perturbácie [5].

3.1.3 Optimalizácia simulovaných prostredí

Na rozdiel od randomizácie, optimalizácia simulovaných prostredí je technika, v ktorej je cieľom lepšie emulovať fyzickú realitu. Samozrejme, aby bolo možné simuláciu optimalizovať, je vyžadované väčšie množstvo dát z reálneho sveta. Tým pádom je táto metóda špecifická pre každého robota a aplikáciu [3]. Doteraz spomenuté techniky redukovania rozdielu v realite spočívali v tom, že v simulácii naučené ovládače bolo možné priamo a jednoducho previezť do reálneho robota.

3.1.4 Transferové učenie

Na problém rozdielu v realite sa z iného uhla pozerá technika nazývaná transferové učenie (angl. transfer learning). Oproti technikám, ktoré sa snažia riešiť rozdiel v realite, cieľom transferového učenia je vyhnúť sa výskytu tohto rozdielu prostredníctvom dvoch súbežných alebo po sebe idúcich fáz tréningu. Prvá tréningová fáza prebieha v simulácii, druhá vo fyzickej realite. Tento spôsob vieme použiť na klasické učenie posilňovaním, teda učenie pomocou metódy pokus a omyl. Poznatky naučené v simulácii vieme využiť v nasledujúcej fáze tréningu v realite, a aj pri neúspešnom pokuse je možné recyklovať nejaké množstvo týchto poznatkov. Tým vieme výrazne znížiť čas potrebný na tréning. Transferové učenie sa delí podľa viacerých vlastností. Jedna z nich je smer prenosu, jednosmerné smeruje len zo simulácie do fyzickej reality, zatiaľ čo obojsmerné prenáša informácie aj opačným smerom. S tým súvisí ďalšia vlastnosť a to počet prenosov, ten môže byť jeden, čo je charakteristické pre jednosmerné transferové učenie, alebo môže prenos nastať viac krát, čo sa využíva pri obojsmernom type. Treťou dôležitou vlastnosťou je druh informácií, ktoré sú prenášané, môže ísť o určité parametre alebo aj o celé stratégie. Na redukcii rozdielu v realite je špeciálne praktický obojsmerný typ transferového učenia, pri ktorom si simulátor prevezme dáta z reálneho sveta a tým vie realistickejšie prispôsobiť prostredie simulácie k reálnemu svetu. Týmto zredukuje odlišnosti medzi virtuálnou a fyzickou realitou, a teda aj samotný rozdiel v realite [5].

3.1.5 Použitie viacerých simulátorov súčasne

Poslednou metódou, ktorú spomenieme je použitie viacerých simulátorov na prekonanie skreslenia jedného simulátoru. Simulátory sú odlišné v rôznych vlastnostiach, preto je možné využiť jeden ako hlavný simulátor a potom na úspešné ovládače použiť

druhý simulátor, ktorý odstráni ovládače, využívané na nerealistické scenáre. Rovnako je možnosťou využiť niekoľko fyzikálnych motorov namiesto jedného [3]. Napriek veľkému počtu rôznych techník na riešenie problému rozdielu v realite, je tento problém do veľkej miery stále prítomný. Dokonca ani nevieme určiť, ktorá z týchto techník je najefektívnejšia. Ako bolo už vyššie spomínané, potrebovali by sme nejaký spôsob ako kvantifikovať rozdiel v realite.

3.2 Nové metódy

3.2.1 Ladenie simulátoru

Collins a kol. [1] ukázali, že rozdiel v realite sa dá do veľkej miery redukovať prostredníctvom ladenia simulátoru, teda úpravy určitých parametrov podľa špecifického scenáru a robota. Táto technika obsahuje štyri fázy učenia. Prvou je kolekcia dát z reálneho sveta. Druhou fázou je samotná simulácia činnosti robota. V testovaní použili simulátory Py-Bullet a V-Rep, obidva používajú fyzikálny motor Bullet. Tieto simulátory ponúkajú širokú škálu parametrov, ktoré je možné nastavovať, preto sú vhodné pre túto metódu. Treťou fázou je selekcia parametrov simulátora. Štvrtou a poslednou fázou je optimalizácia simulátoru nastavením parametrov vybraných v tretej fáze. Na optimalizáciu využili algoritmus nazývaný diferenciálna evolúcia implementovaný v optimalizačnom module SciPy, pretože umožňuje optimalizovať veľké množstvo parametrov súčasne. Pri testovaní optimalizácie zistili, že najdôležitejšie parametre sú časový krok simulácie, laterálne trenie objektov a kĺbová rýchlosť. Táto technika vykazuje dobré výsledky aj pri interakcii s objektami.

3.2.2 Diferencovateľná fyzika

S narastajúcim počtom dostupných knižníc pre automatickú diferenciáciu prichádzajú do popredia simulátory, ktoré využívajú diferencovateľné fyzikálne motory. Tieto simulátory majú v robotike veľký potenciál, pretože sľubujú optimalizáciu simulačných parametrov, riadiacich vstupov a morfológie agenta [2].

Collins a kol. [2] navrhli novú revolučnú metódu na riešenie problému rozdielu v realite, ktorá využíva tieto knižnice. Nazvali ju RealityGrad. Pozostáva zo štyroch fáz, podobne ako ladenie simulátoru (časť 3.2.1). Prvou fázou je kolekcia optimálnych trajektórií v simulácii, ktorá bude použitá ako množina dát v nasledujúcej fáze. Druhou fázou je tréning stratégií a treťou je následné zavedenie týchto stratégií, teda ich použitie na robotovi v reálnom svete. Štvrtou a najdôležitejšou fázou je systémová identifikácia. V nej sa dáta nazbierané v tretej fáze využívajú na zlepšenie presnosti simulátoru pomocou knižnice pre automatickú diferenciáciu. Tieto knižnice umožňujú

túto optimalizáciu simulátoru spraviť časovo efektívne, vyžadujú ale väčšie množstvo pamäte RAM. Pre ešte presnejšie výsledky sa tieto štyri spomenuté fázy vykonajú niekoľko krát po sebe, s tým, že prvá iterácia berie dáta z reálneho sveta a každá ďalšia berie už optimalizované dáta predošlej iterácie. Tým je zaručené neustále spresňovanie simulátoru, a teda aj násobné zredukovanie rozdielu v realite.

Literatúra

- [1] Jack Collins, Ross Brown, Jurgen Leitner, and David Howard. Traversing the reality gap via simulator tuning, 2020. [arXiv:2003.01369](#).
- [2] Jack Collins, Ross Brown, Jürgen Leitner, and David Howard. Follow the gradient: Crossing the reality gap using differentiable physics (realitygrad), 2021. [arXiv:2109.04674](#).
- [3] Jack Collins, David Howard, and Jürgen Leitner. Quantifying the reality gap in robotic manipulation tasks, 2018. [arXiv:1811.01484](#).
- [4] Quentin Le Lidec, Wilson Jallet, Louis Montaut, Ivan Laptev, Cordelia Schmid, and Justin Carpentier. Contact models in robotics: a comparative analysis, 2023. [arXiv:2304.06372](#).
- [5] Erica Salvato, Gianfranco Fenu, Eric Medvet, and Felice Andrea Pellegrino. Crossing the reality gap: A survey on sim-to-real transferability of robot controllers in reinforcement learning. *IEEE Access*, 9:153171–153187, 2021. [doi:10.1109/ACCESS.2021.3126658](#).