

Sledování pohybujících se objektů prostřednictvím videokamery

Jan Krejčí¹

1 Úvod

Sledování pohybujících se objektů je úloha, jejímž cílem je odhadovat počet a polohy (popř. rychlosti, trajektorie, druhy) objektů využitím dat ze senzorů. Je-li senzorem videokamera, hovoříme o vizuálním sledování. Jako příklad využití vizuálního sledování uvedme sledování okolí pozemní komunikace autonomním vozidlem za účelem vyvarování se srážky.

Obrazová data jsou doménou především počítačového vidění. Algoritmy vizuálního sledování z této oblasti často využívají metody, jejichž aplikace může být v praktických úlohách výpočetně náročná. Jako řešení se nabízí využít pouze algoritmy zvané vizuální detektory. Ty zpracovávají každý snímek z videa zvlášť, přičemž jejich výstupy jsou typicky množiny obdélníků ohraničující nalezené objekty. Informace o jednotlivých trajektoriích tedy chybí. Navíc může z různých důvodů docházet k chybějícím, nebo chybným detekcím objektů.

Nezávisle na počítačovém vidění jsou v oblasti teorie odhadu vyvíjeny algoritmy pro sledování objektů sekvenčním zpracováním dat. Tyto algoritmy budeme nazývat filtračními. Bývají často využívány v případě, že senzorem je radar/sonar a jejich výhodou je mimo jiné modelování pohybu objektů, což umožňuje řešit problémy s chybějícími nebo chybnými detekcemi. To z nich dělá vhodné kandidáty na využití v úloze vizuálního sledování, kdy jako vstupní data filtru používáme výstupy vizuálního detektoru. V tomto příspěvku popíšeme jak filtrační algoritmy pracují a naznačíme jejich možné využití v úloze vizuálního sledování.

2 Filtrační algoritmy

Pohybující se objekt lze často věrně charakterizovat jeho aktuální polohou a rychlostí, případně dalšími veličinami. Uspořádáme-li tyto veličiny do vektoru, nazýváme ho stavem objektu. V teorii odhadu se vektor stavu považuje za náhodný. Dále předpokládáme, že se stav v čase vyvíjí podle zvoleného modelu pohybu, během čehož senzor generuje nepřesná měření. V těchto měřeních (a modelu pohybu) je obsažena veškerá dostupná informace o stavu. Cílem filtračních algoritmů je tuto informaci extrahovat. V případě, že data zpracováváme sekvenčně, to často znamená rekursivně počítat hustotu pravděpodobnosti stavu, podmíněnou všemi dostupnými měřeními. Výpočet této hustoty se v daném čase uskutečňuje ve dvou krocích. V prvním je využít zvolený model pohybu k predikci stavu do následujícího časového okamžiku, a ve druhém je pomocí Bayesova vztahu využita informace z měření.

V případě, že objektů chceme sledovat více, je možné problém formulovat, i řešit analogicky. Namísto jednoho stavu můžeme modelovat množinu stavů, jejíž mohutnost i prvky jsou uvažovány jako náhodné. Protože z každého objektu může (nemusí) vzniknout měření, můžeme i tato měření modelovat jako množinu. Hledaná hustota pravděpodobnosti je potom funkce konečné množiny stavů, kde i její mohutnost je volnou proměnnou.

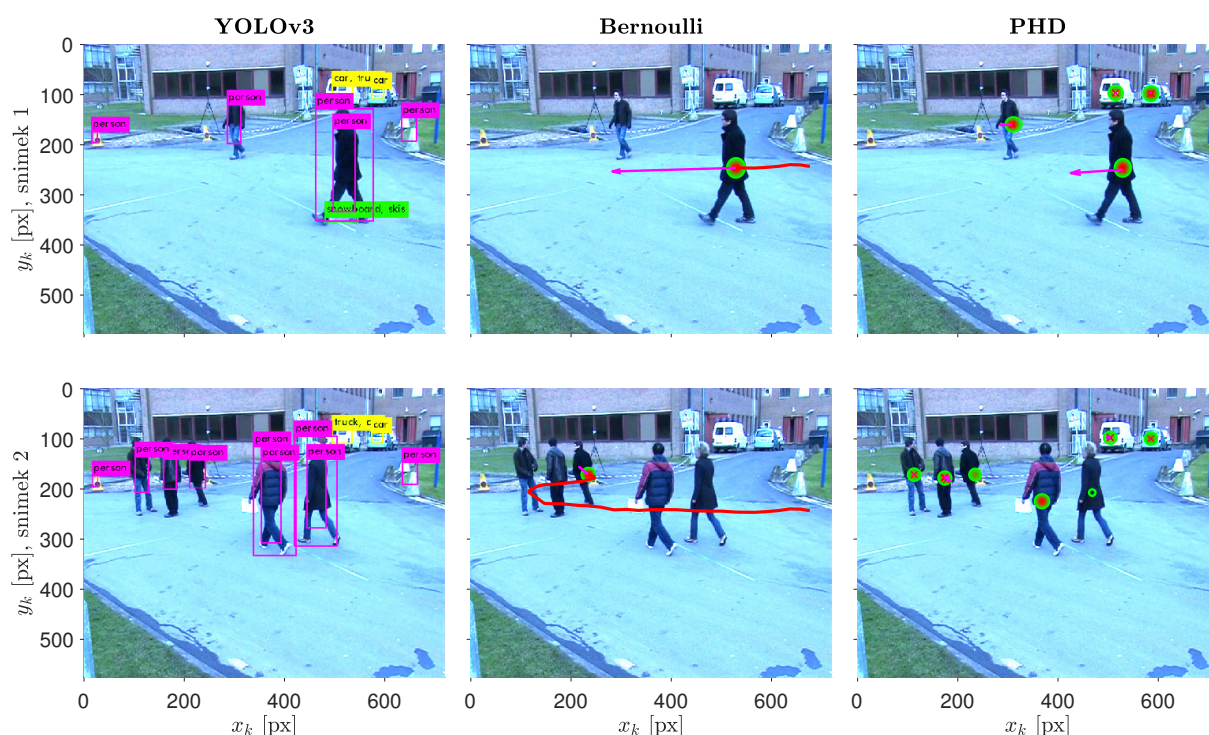
¹ student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, e-mail: jkrejci@students.zcu.cz

Stěžejní filtrační algoritmy jsou shrnuty v knize Streit et al. (2021). Pro jednoduchost rozdělme tyto algoritmy na dvě skupiny podle filosofie modelování objektů. První skupina modeluje objekty zvlášť a vyhodnocuje hypotézy o přiřazení měření k objektům. Příkladem je Bernoulliho filtr, který předpokládá jeden objekt, a odhaduje pravděpodobnost jeho existence. Multi-Bernoulliho filtr pak modeluje sjednocení více takových objektů.

V druhé skupině jsou filtry modelující všechny objekty pomocí jediné funkce stavu. Příkladem je PHD filtr, z anglického *probability hypothesis density*. Ten předpokládá, že počet objektů má Poissonovo rozdělení, jehož střední hodnota je zde odhadovanou veličinou. Tento algoritmus má nízké výpočetní nároky za cenu neposkytování odhadů trajektorií objektů.

3 Aplikace na vizuální data

Využití výše popsaných filtračních algoritmů v úloze vizuálního sledování je ilustrováno na videu z datasetu Ferryman a Shahrokni (2009), dva ukázkové snímky jsou na obrázku 1. Vrstevnice výsledné hustoty pravděpodobnosti jsou znázorněny zeleno-červeně, směr pohybu (rychlost) naznačena fialovou šipkou a trajektorie (pouze Bernoulliho filtr) červenou křivkou. Jako vizuální detektor byl využit algoritmus *you only look once* verze 3 (YOLOv3), přičemž jako měření pro filtrační algoritmy byly využity středy nalezených obdélníků. Pro časový vývoj stavů byl předpokládán model s téměř konstantní rychlostí.



Obrázek 1: Ukázka výsledků z videa *View_006_T12-T34* pro Bernoulliho a PHD filtr.

Literatura

Streit, R., Angle, R. B. a Efe, M. (2021) *Analytic Combinatorics for Multiple Object Tracking*, Springer Nature Switzerland AG

Ferryman, J. a Shahrokni, A. (2009) *PETS2009: Dataset and challenge* in 2009 Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, pp. 1-6