

SPOJENÍ DYNAMICKÝCH DAT ROZPOZNANÝCH UMĚLOU INTELIGENCÍ S 3D MODELEM MĚSTA PLZNĚ

Pavel HÁJEK¹, Lukáš PICEK², Michal KEPKA¹, Tomáš TICHÝ³, Radek FIALA¹, Karel JANEČKA¹

¹ Katedra geomatiky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni,
Univerzitní 2732/8, 301 00, Plzeň, ČR

gorin@kgm.zcu.cz, mkepka@kgm.zcu.cz, fialar@kgm.zcu.cz, kjanecka@kgm.zcu.cz

² Katedra kybernetiky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni,
Univerzitní 2732/8, 301 00, Plzeň, ČR

picekl@kky.zcu.cz

³ Katedra dopravní telematiky, Fakulta dopravní, České vysoké učení technické v Praze,
Konviktská 20, Praha, ČR

tomas.tichy@cvut.cz

DOI 10.31490/9788024847955-3

Abstrakt

Rychlý rozvoj konceptu digitálních dvojčat v různých průmyslových odvětvích klade vysoké nároky nejen na odborné znalosti, ale také na dostupnost, kvalitu a integraci nových typů dat do stávajících modelů. Jednou z klíčových výzev je začlenění dynamických prvků, tj. objektů měnících svou polohu v čase a prostoru, do těchto modelů. Tento příspěvek představuje komplexní workflow pro propojení dynamických dopravních dat, získaných pomocí umělé inteligence, se statickým 3D modelem města Plzně. Postup zahrnuje sběr dat, tj. detekci, rozpoznávání a sledování objektů a jejich následnou vizualizaci na virtuálním 3D glóbu. Dynamické objekty, jako jsou chodci či vozidla, jsou detekovány pomocí neuronových sítí a umožňují zpracování ze dvou zdrojů: statické kamery snímající vybranou oblast a kamery umístěné na pohybující se tramvaji. Pro správné propojení dynamických objektů s 3D modelem města byla provedena projektivní transformace obrazových souřadnic na geografické souřadnice cílového systému. Transformovaná data, zahrnující časové a prostorové aspekty, byla následně vizualizována ve webové mapové aplikaci postavené na platformě CesiumJS. Výsledná aplikace zobrazuje dynamické objekty, jako jsou chodci, automobily, motocykly apod. v kombinaci se statickými 3D modely okolní zástavby. Tento přístup otevírá nové možnosti pro analýzu časoprostorové mobility ve městech, nástroje pro efektivní detekci a reprezentaci zájmových objektů a přístupnou vizualizaci dynamických prvků prostřednictvím webových aplikací.

Abstract

The rapid development of the digital twin concept across various industrial sectors imposes significant demands on specialised knowledge and the availability, quality, and integration of new data types into existing models. One of the key challenges lies in incorporating dynamic elements, i.e., objects that change their position over time and space, into these models. This paper presents a comprehensive workflow for integrating dynamic traffic data obtained using artificial intelligence with a static 3D model of the city of Pilsen. The workflow includes data collection – detection, recognition, and tracking of objects – and their subsequent visualisation on a virtual 3D globe. Dynamic objects such as pedestrians and vehicles are detected using neural networks and processed from two sources: stationary cameras monitoring a selected area and cameras mounted on a moving tram. For proper integration of dynamic objects with the 3D city model, a projective transformation of image coordinates to geographic coordinates of the target system was performed. The transformed data, incorporating temporal and spatial aspects, were then visualised in a web mapping application built on the CesiumJS platform. The resulting application displays dynamic objects, such as pedestrians, cars, motorcycles, etc., in combination with static 3D models of surrounding buildings. This approach opens new possibilities for analysing spatiotemporal mobility in cities, tools for efficient detection and representation of objects of interest, and accessible visualisation of dynamic elements through web applications.

Klíčová slova: dynamický objekt; rozpoznávání; umělá inteligence; 3D model města; mobilita; vizualizace; webová mapová aplikace

Keywords: dynamic object; recognition; artificial intelligence; 3D city model; mobility; visualization; web map application

ÚVOD

Začlenění dynamických prvků do 3D modelů měst se pro urbanisty, architekty a tvůrce politik již stává prakticky nezbytností. Tradiční statické 3D modely měst jsou sice neocenitelné pro vizualizaci a prostorovou analýzu, ale často je potřeba zachycovat i vyvíjející se povahu městského prostředí. Dynamické prvky, jako je dopravní tok, pohyb chodců a podmínky prostředí, jsou zásadními složkami, které utvářejí funkčnost a obyvatelnost měst. Proto existuje naléhavá poptávka po rozšíření stávajících 3D modelů měst o data a dynamické simulace v reálném čase, které by umožnily komplexnější pochopení dynamiky měst. Tento příspěvek se zaměřuje na rozpoznávání a integraci dynamických prvků do stávajících 3D modelů měst optikou digitálních dvojčat, přičemž se zaměřuje zejména na dynamiku dopravy a chodců zachycené na kamerových záběrech a rozpoznané pomocí umělé inteligence, spolu s vizualizací polohy rozpoznávaných objektů ve virtuální 3D scéně.

SBĚR A ZPRACOVÁNÍ DAT PRO DYNAMICKOU EXTRAKCI OBJEKTŮ

Plzeň, jako významné dopravní centrum kraje, se vyznačuje klíčovou rolí veřejné dopravy, která zahrnuje jak vnitřní městský systém dopravy, tak vnější propojení s širším regionem. Jak uvádí Tichý a kol. [1], digitalizace dat v takovém dopravním systému může být realizována různými způsoby a současně lze tato data využít nejen pro statistické sledování a hodnocení, ale také pro různé platformy. V rámci projektu „Vytvoření dynamického digitálního modelu ulice pro potřeby autonomního řízení v Plzni“ (Výzkumný projekt “DiDYMOS”, viz kapitola Poděkování) byl jeden z hlavních cílů zaměřen na integraci dynamických dat do stávajícího statického 3D modelu města. Zatímco data modelu města poskytuje Správa informačních technologií města Plzně spravující městská prostorová data, dynamická data jsou získávána přímo z reálného prostředí skrze senzory, v tomto případě konkrétně statické kamery.

Pro zpracování dynamických prvků byla vytvořena efektivní linka založená na moderních algoritmech strojového učení. Tyto algoritmy umožňují detekci a sledování objektů z kamerových záznamů, přičemž detekované objekty jsou následně vizualizovány ve virtuálním 3D prostředí. Proces vizualizace zahrnuje perspektivní transformaci souřadnic polohy a časovou synchronizaci detekovaných objektů, díky čemuž lze jejich pohyb přesně zaznamenat na 3D glóbu. Další části textu podrobně popisují jednotlivé kroky tohoto procesu.

Pro návrh nejvhodnějšího a robustního řešení pro extrakci dynamických objektů jsme shromáždili datový soubor obsahující 24 hodin záznamů. Tyto záznamy byly pořízeny statickou kamerou za různých povětrnostních a světelných podmínek, viz obrázek 1. Všechny záznamy mají rozlišení 3840 × 2160 px a byly snímány rychlostí 25 snímků za sekundu.



Obr. 1. Vybraná oblast zájmu pro extrakci dynamických prvků z pohledu statické kamery – v levé části za dobrých povětrnostních podmínek, v pravé části za deště.

Výsledky kalibračního procesu jsou znázorněny na obrázku 2. Zelené úsečky reprezentují výšky kalibračních objektů (například dopravních značek) v jejich příslušných polohách. Červené úsečky znázorňují referenční výšku 2 metry od povrchu země, která je rovnoměrně rozmístěna v prostoru zachycené scény.



Obr. 2. Výsledek kalibrace založený na zelených liniích představujících výšky kalibračních objektů.

VÝSLEDKY

V rámci zmíněného výzkumného projektu DiDYMOS byl vyvinut software pro rozpoznávání statických a dynamických objektů z videosekvencí. Tento software¹ umožňuje:

- Detekci objektů ve scéně, jako jsou dopravní značení, koleje, automobily nebo osoby.
- Rozlišení mezi statickými a dynamickými objekty.
- Sledování dynamických objektů a jejich rozdělení podle unikátních identit, což umožňuje další analýzu jejich trajektorií.

Pro využití stacionární kamery byl vyvinut algoritmus pro rozpoznání a sledování pohybujících se objektů, jako jsou chodci, automobily, motocykly a tramvaje. Pro detekci a následné rozpoznání objektů byly využity neuronové sítě typu YOLOv8, které vyvinul Jocher a kol. [2]. Zatímco sledování jejich pohybu v čase zajišťuje algoritmus ByteTrack, který popsali Zhang a kol. [3]. Modul obsahuje také funkci pro kalibraci scény, která využívá měření velikostí statických objektů ve scéně (viz obrázek 2). Kalibrační model předpokládá, že sledovaný zemský povrch je rovina, což umožňuje:

- Odhad velikosti objektů v metrech.
- Robustnější detekci objektů v situacích, kdy:
 - velikost objektů je nestabilní,
 - objekty jsou částečně překryty jinými,
 - nebo opouštějí snímanou scénu.

Kalibrace rovněž umožňuje mapovat trajektorie objektů na mapové podklady a tím identifikovat kritické oblasti (například přejezdy či přechody). Tímto způsobem lze detekovat přítomnost nežádoucích objektů – například chodce v kolejišti, který je označen modře v červeně vyznačené rizikové oblasti (viz obrázek 3 vpravo). Pro zlepšení přesnosti sledování trajektorií jednotlivých objektů je využit Kalmanův filtr, který vyhlazuje výsledné trajektorie a zvyšuje jejich spolehlivost. Na obrázku 3 jsou zobrazeny výsledky sledování objektů ve scéně. Levá část obrázku ukazuje objekty a jejich stopy v pohledu z kamery, zatímco pravá část zobrazuje jejich polohy v rovině povrchu při pohledu shora.

¹ Dostupný na <https://kky-sw.zcu.cz/cs/sw/DynObT>. Další informace na <https://sites.google.com/view/didymos-tacr/results/scene-objects-detection>



Obr. 3. Ukázka sledovaných objektů zájmu – z pohledu kamery (vlevo) a z pohledu shora (vpravo).

Pro spojení dynamických prvků s 3D modelem města je nezbytné používat standardní prostorový referenční systém, v našem případě systém zeměpisných souřadnic. Dynamické prvky, které byly rozpoznány na snímcích, mají však pouze obrazové souřadnice. K jejich transformaci na zeměpisné souřadnice jsme pro testovací účely použili jednoduchou 2D homografii (perspektivní transformaci). Tento přístup předpokládá, že dynamické prvky leží na předem definované horizontální rovině v reálném světě. Ačkoli není tento předpoklad zcela přesný, pro naše potřeby poskytuje dostatečně uspokojivé výsledky.

Parametry transformace byly odhadnuty z odpovídajících dvojic obrazových a zeměpisných souřadnic metodou nejmenších čtverců (viz zelené body na obrázku 4). Následně se tato transformace aplikovala na všechny dynamické prvky v různých časových okamžicích, čímž se jejich obrazové souřadnice převedly na zeměpisné. Poloha kamery (červená tečka), odpovídající dvojice obrazových a zeměpisných souřadnic (zelené body) a přibližná snímaná oblast (červený čtyřúhelník) jsou znázorněny na obrázku 4.



Obr. 4. Poloha kamery (červená tečka), oblast snímání (červená oblast) a transformační body (zelené tečky).

Pro vizualizaci dynamických prvků byla použita knihovna CesiumJS [4], která je moderním nástrojem pro 3D vizualizaci různorodých dat. CesiumJS umožňuje snadné začlenění časově dynamických prvků prostřednictvím svého interního formátu Cesium Language (CZML) [5]. Tento formát umožňuje přesně definovat změny hodnot dynamických prvků během daného časového období. Díky tomu umožňuje přesné zobrazení dynamických bodů ve scéně, včetně jejich časové polohy a synchronizaci vizualizace s výchozím časováním kamery, což zajišťuje konzistenci při zobrazení dynamických prvků ve vztahu ke statickému obsahu. Ukázka vizualizace v knihovně CesiumJS je na obrázku 5. V rámci ní se zobrazuje kombinace statické podkladové mapové vrstvy (ortofoto ČR jako WMS) se statickým 3D modelem města (bílé budovy ve formátu GLB) a dynamickými bodovými prvky vyjádřenými kartografickými symboly podle typu prvku – tramvaj, auto, motocykl či osoba, vizte detail ve spodní pravé části obrázku 5). Vizualizační webová aplikace v CesiumJS je k dispozici na osobních univerzitních stránkách autora².

² https://home.zcu.cz/~mkepka/didymos/dt2/history_cesium.html



Obr. 5. Ukázka z vizualizační webové aplikace CesiumJS zobrazující statické prvky budov a dynamické prvky automobilů, motocyklu, osoby a tramvaj vyskytujících se v daném čase v zájmové oblasti.

DISKUZE A ZÁVĚR

Tento příspěvek se zaměřil na integraci dynamických prvků do 3D modelu města a jejich vizualizaci. Pokrývá oblasti 3D modelování měst, 3D vizualizace, časoprostorové GIS a geoprostorová data – statická i dynamická. Tyto přístupy směřují k rozvoji městských digitálních dvojčat, která mohou výrazně přispět k efektivnímu řízení a plánování městské infrastruktury.

Pro získání dynamických rysů, tedy informací o pohybujiících se objektech, jako jsou chodci a vozidla, byla vyvinuta technologická linka založená na pokročilých algoritmech strojového učení, která zahrnovala:

1. Detekci a sledování objektů pomocí modelů YOLOv8 a ByteTrack.
2. Kalibraci kamery a transformaci obrazových souřadnic na zeměpisné.
3. Integraci dynamických objektů do statického 3D modelu města.
4. Vizualizaci dynamických prvků prostřednictvím knihovny CesiumJS s využitím formátu CZML.

Tento proces umožnil časoprostorovou vizualizaci dynamických objektů, která zobrazuje jejich polohu a časový výskyt v městském prostředí. Výsledky byly prezentovány ve webové aplikaci, což umožňuje snadnou dostupnost a flexibilní interakci s daty. Navržená technologická linka má však určitá omezení, která otevírají prostor pro budoucí výzkum a vývoj jako např. použití jedné statické kamery, převod souřadnic pouze ve 2D rovině, rozpoznávání jen vybraných kategorií objektů (osoba, silniční vozidla, tramvaj) či jednoduchá kartografická vizualizace.

I přes tato omezení poskytuje řešení robustní, dostupnou a replikovatelnou platformu, na které lze stavět při dalším vývoji. Klíčové body, jako je integrace dynamických prvků s 3D modely a jejich časoprostorová vizualizace, představují silný základ pro rozvoj pokročilých městských digitálních dvojčat.

PODĚKOVÁNÍ

Tento příspěvek byl podpořen ze státního rozpočtu Technologickou agenturou České republiky a Ministerstvem dopravy České republiky na základě smlouvy programu TRANSPORT 2020+ č. CK03000179 „Vytvoření dynamického digitálního modelu ulice pro potřeby autonomního řízení v Plzni“ – DiDYMOS, <https://starfos.tacr.cz/cs/projekty/CK03000179>.

LITERATURA

Tichý, T., Brož, J. Růžička, J., Hejčiarová, E., Bělinová, Z., Jíšová, J., Filip, J., Navrátilová, K., Ivasienko, P., Šmerda, T. (2023) "Information modeling and Smart Approaches at the Interface of Road and Rail Transport," 2023 Smart City Symposium Prague (SCSP), Prague, Czech Republic, pp. 1-8, doi: 10.1109/SCSP58044.2023.10146231.

Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8 (Version 8.0.0). Retrieved from <https://github.com/ultralytics/ultralytics>.

Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W. & Wang, X. (2022) Bytetrack: Multi-object tracking by associating every detection box. In European conference on computer vision (pp. 1-21). Cham: Springer Nature Switzerland. <https://doi.org/10.48550/arXiv.2110.06864>.

Cesium GS, Inc. (2023) CesiumJS platform, <https://cesium.com/platform/cesiumjs/>. [cit. 19. 12. 2024]

Hunter, S. (2016) Cesium Language (CZML) Guide, <https://github.com/AnalyticalGraphicsInc/czml-writer/wiki/CZML-Guide>. [cit. 19. 12. 2024]