

Kooperatív vizuális SLAM rendszer járműalapú környezetérzékelésre

Fodor Károly¹ – Rövid András²

¹Doktornadusz, Budapesti Műszaki és Gazdaságtudományi Egyetem Közlekedés és Járóműmérnöki Kar, Gépjárműtechnológiai Tanszék

²Tudományos főmunkatárs, Budapesti Műszaki és Gazdaságtudományi Egyetem Közlekedés és Járóműmérnöki Kar, Gépjárműtechnológiai Tanszék

e-mail: fodor.karoly@kjk.bme.hu, rovid.andras@kjk.bme.hu

Kézirat benyújtva: 2026.01.10.

Kézirat elfogadva: 2026.03.30.



© Fodor K., Rövid A.

Absztrakt

Az önvezető járművek számára kulcsfontosságú a környezet pontos felmérése és a megbízható lokalizáció. Ehhez egy kooperatív vizuális SLAM rendszert ismertetünk, ahol az egyik jármű monokuláris kamerával térképet készít, és ezt a térképet GNSS referencia koordinátákkal globális méretarányra kalibrálja. A második jármű a létrehozott már globális koordináta-rendszerben értelmezett térképet felhasználva vizuálisan lokalizálja magát. A bemutatott módszer kiértékeléshez egy saját készítésű adatszettet használtunk. Az eredmények alapján megállapítható, hogy a módszer növeli a helymeghatározás pontosságát, és lehetővé teszi a térképek újrafelhasználását a járművek között.

Kulcsszavak: lokalizáció, GNSS-integráció, kooperatív vSLAM, önvezető járművek

DOI: <https://doi.org/10.24228/KTSZ.2026.3.4>

A könnyebb értelmezhetőség és áttekinthetőség érdekében a cikk szövegében alkalmazott rövidítések magyarázatait az 1. táblázat tartalmazza.

1. BEVEZETÉS

A modern, intelligens járművek számára elengedhetetlen a környezet pontos érzékelése és a jármű pozíciójának megbízható meghatározása. Ehhez gyakran használnak fedélzeti szenzorokat, köztük kamerákat, lidarokat, GNSS-t és/vagy inerciális mérőegységeket. A vizuális SLAM algoritmusok lehetővé teszik, hogy egy jármű a kameraképek alapján egyszerre térképet építsen és meghatározza saját helyzetét ezen a térképen.

Az ORB-SLAM3 egy korszerű, nyílt forráskódú vizuális SLAM algoritmus, amely monokuláris

kamerával is képes valós idejű térképezésre és lokalizációra (Campos et al., 2021). A hagyományos monokuláris SLAM rendszerek ugyanakkor nem tudják meghatározni a térkép abszolút méretét. A monokuláris pozíció gráf hibája út-időben folyamatosan növekszik, így a térkép egyre torzultabbá válik. Ennek kiküszöbölésére és a globális helymeghatározás biztosítására integrálható a GNSS-technológia a vizuális SLAM folyamataiba.

Jelen cikk egy olyan kooperatív vSLAM rendszert mutat be járművek számára, amelyben az első jármű kamerás érzékeléssel térképet épít, és a GNSS mérések felhasználásával biztosítja

Rövidítés	Angol megnevezés	Magyar jelentés/leírás
SLAM	Simultaneous Localization and Mapping	Egyidejű helymeghatározás és térképezés; a rendszer egyszerre becsüli saját helyzetét és építi fel a környezet térképét.
GNSS	Global Navigation Satellite System	Globális navigációs műholdrendszer; a műholdas helymeghatározó rendszerek összefoglaló neve.
IMU	Inertial Measurement Unit	Inerciális mérőegység; gyorsulást és szögsebességet mérő szenzoregység.
BA	Bundle Adjustment	Kötegelt kiegyenlítés; kamerapózok és térképpontok együttes optimalizálása a vetítési hibák csökkentésére.
LiDAR	Light Detection and Ranging	Lézeralapú távolságmérő rendszer; lézersugarak segítségével állít elő távolság- és térbeli információt.
PTP	Precision Time Protocol	Precíziós időszinkronizációs protokoll; különböző szenzorok időbelyegeinek pontos összehangolására szolgál.
SE (3)	Special Euclidean Group in 3D	A háromdimenziós merevtest-mozgások matematikai csoportja; eltolást és forgatást tartalmaz, skálázást nem.
Sim(3)	Similarity Transformation Group in 3D	Háromdimenziós hasonlósági transzformációk csoportja; eltolást, forgatást és skálázást is tartalmaz.
UTM	Universal Transverse Mercator	Egy vetületi koordinátarendszer, amely a földrajzi pozíciók síkbeli, méterben értelmezhető ábrázolását teszi lehetővé.
UGBA	UTM-based Global Bundle Adjustment	UTM-alapú globális kötegelt kiegyenlítés; olyan végső optimalizáció, amely a vizuális és UTM-alapú GNSS információkat együtt használja.
APE	Absolute Pose Error	Abszolút pózhiba; a becsült és referencia pálya közötti abszolút helyzeti eltérés mérőszáma.
ORB	Oriented FAST and Rotated BRIEF	Képelem-detektor és leíró (descriptor) algoritmus, amelyet a kamera képen található jellegzetes pontok (sarkok, élek) gyors felismerésére és követésére használnak.
vSLAM	Visual Simultaneous Localization and Mapping	Vizuális egyidejű helymeghatározás és térképezés.
LD-SLAM	Laser/LiDAR-based Simultaneous Localization and Mapping	Lézer alapú egyidejű helymeghatározás és térképezés.
RGB-D	Red, Green, Blue – Depth	A számítógépes látásban és képkalkotásban használt kifejezés, amely a hagyományos színes képkalkotást kombinálja a távolságadatokkal.

1. táblázat: A cikkben alkalmazott rövidítések magyarázatai

a térkép valós méretét és globális pozícióit. Az általa létrehozott pontos, abszolút térképet egy második jármű felhasználhatja, és saját kamerája alapján lokalizálni tudja magát a megosztott térképen belül. Ezzel a közös rendszerrel a járművek megoszthatják egymás közt a környezeti információkat, javul a helymeghatározás pontossága, és csökken az egyes járművek által végzett párhuzamos adatgyűjtés igénye. A következőkben megvizsgáljuk a kapcsolódó irodalmat, részletesen ismertetjük a rendszer felépítését, a GNSS-integráció módját, továbbá kitérünk a megoldás előnyeire és a lehetséges további szenzor bővítési irányokra is. A módszert ezután egy saját készítésű adatszettben kiértékeljük.

2. KAPCSOLÓDÓ KUTATÁSOK

A vizuális SLAM algoritmusok fejlődése az elmúlt évtizedben jelentős előrelépéseket hozott a környezetérzékelés és helymeghatározás terén. Az ORB-SLAM család algoritmusai – különösen az ORB-SLAM2 és a 2021-ben bemutatott ORB-SLAM3 – már képesek monokuláris,

sztereó és RGB-D kameraképek feldolgozására, valamint inerciális szenzorral történő kiterjesztésre is (Mur-Artal et al. 2015; Mur-Artal & Tardós, 2017; Campos et al., 2021). Az ORB-SLAM3 egyik fő előnye, hogy szálanként elkülöníti a képkövetést, a lokális térképezést és a globális optimalizálást, így nagy térben is képes valós idejű működésre. Továbbá képes több térkép együttes kezelésére. A térképezést aktív és inaktív térképekre osztja. A módszer képes térképek tárolására és újrafelhasználására (Campos et al., 2021).

A monokuláris vizuális SLAM rendszerek gyengesége ugyanakkor a globális referencia hiánya: a térképek relatív koordináta-rendszerben épülnek fel, így a valós helyzet és skála ismeretlen. Tetszőleges skála felvételével történik az inicializálás, azonban ez a felvett skála fokozatosan torzul. Ennek kiküszöbölésére több kutatás is foglalkozott GNSS adatok SLAM-be történő integrálásával. Kiss-Illés et al. (2019) a GPS-SLAM rendszerben a monokuláris SLAM által becsült lokális pályát műholdas mérésekkel támogatták annak érdekében, hogy stabilabb globális pozícióbecslést érjenek el. A módszer elsősorban

nem globális optimalizációra épül, hanem a GNSS és IMU adatok felhasználásával javítja a vizuális követés póz-előrejelzését. A szerzők kimutatták, hogy már ritkább és zajosabb GNSS mérések is elegendőek lehetnek a globális sodródás csökkentésére és a pálya abszolút koordináta-rendszerhez való rögzítésére, különösen alacsony képkockaszámú felvételek esetén.

Chen et al. (2018) a GVORB rendszerben az ORB-SLAM lokális és globális kötegelt igazítási (továbbiakban BA) eljárás közé egy köztes méretű optimalizációs lépést illesztettek, amely már GNSS-alapú, abszolút pozíció méréseket is tartalmaz. Ez a köztes optimalizáció nagyobb kiterjedésű, mint a lokális BA, ugyanakkor számításiigénye kisebb a teljes globális BA-nál, így lehetővé teszi a globális sodródás folyamatos, valós idejű csökkentését. A rendszer többszálú architektúrában működik, ahol a GNSS mérésekkel kiegészített optimalizáció elkülönített szálakon fut, megőrizve a vizuális SLAM lokális stabilitását.

Engel et al. (2018) az LD-SLAM rendszerben a nagy léptékű direkt SLAM skálázhatóságát lokális térképrészletek alkalmazásával biztosítják. Az egyes térképrészletek közötti kapcsolatot lokálisan becsült hasonlósági transzformációk írják le, ami lehetővé teszi a valós idejű követést és térképezést, ugyanakkor globális inkonzisztenciákat és fragmentációt eredményezhet hosszú trajektóriák esetén. A rendszer ezért időszakosan egy globális optimalizációs lépést futtat, amely a lokálisan becsült hasonlósági transzformáció-kapcsolatok összehangolásával helyreállítja a globális konzisztenciát. Ez a globális optimalizáció aszinkron, háttér folyamatként működik, így nem blokkolja a valós idejű működést, de nem tekinthető folyamatosan valós idejűnek.

A kooperatív SLAM rendszerek esetében további kihívás a térképek megosztása és az együttműködés: Wang et al. (2022) megállapították, hogy a térképek újrafelhasználása más ágens lokalizációjára jelentős számítási kapacitást takaríthat meg, miközben a GNSS-re épülő térképek közös referencia-rendszerként is szolgálhatnak. Ezen alapelvek mentén jelen tanulmány célja egy olyan vSLAM rendszer megalkotása, amely valós idejű működés mellett képes GNSS-pontosságú térképet előállítani és azt kooperatív módon lokalizációs célra újrafelhasználni.

2. 1. Alternatív érzékelők a skálázás támogatására

A bemutatott rendszerben a GNSS szolgáltatja az abszolút skálát és a globális referenciát. Ugyanakkor léteznek más szenzorok és mérőrendszerek is, amelyek segítségével a vizuális SLAM skálázása és pontossága tovább javítható, illetve GNSS hiányában is biztosítható a helyes méretarány (Scaramuzza és Fraundorfer, 2011; Cadena et al., 2016).

Az alábbiakban a legfontosabb alternatív megoldásokat foglaljuk össze:

Kerékodometria használata esetén a jármű kerékeinek forgásából származtatott megtett út értéke alkalmas lehet a vizuális SLAM méretarányának kalibrálására. A kamerából becsült elmozdulás összevethető a kerékodometria által mért távolsággal, amelyből abszolút skálafaktor határozható meg (Geiger et al., 2012). A módszer hátránya, hogy a kerékmegecsúszás és a talajviszonyok jelentős hibát okozhatnak, ezért önmagában nem tekinthető megbízhatónak. Vizuális SLAM-mel kombinálva azonban hatékony kiegészítő információt biztosíthat (Zhang és Singh, 2014). A giroszkópból és gyorsulásmérőből álló inerciális mérőegység (IMU) folyamatosan méri a jármű forgásait és gyorsulásait. A vizuális SLAM-mel integrálva javítja a pályabecslés stabilitását, és a gravitációs gyorsulás ismert értékének felhasználásával a rendszer közvetlenül képes meghatározni a skálát (Leutenegger et al., 2015). Az IMU különösen hasznos olyan rövid időszakokban, amikor a vizuális információ gyenge vagy hiányos, például gyors mozgás vagy elmosódás esetén (Qin et al., 2018). A lézeralapú távolságmérők közvetlen, metrikus távolságmérést biztosítanak, így a LiDAR-alapú SLAM rendszerek alapvetően nem szenvednek skálabizonytalanságtól (Zhang és Singh, 2014). A kamerás és LiDAR-alapú érzékelés kombinálásával a vizuális SLAM robusztussága és pontossága tovább növelhető, különösen strukturálatlan vagy rosszul megvilágított környezetekben (Behley és Stachniss, 2018).

Az alternatív szenzorok bevonásával a kooperatív vizuális SLAM rendszer rugalmasan adaptálható különböző környezeti és üzemeltetési feltételekhez. GNSS-szel nem lefedett területeken – például alagutakban vagy beltéri környezetben – az inerciális és odometriai adatok

meghatározóvá válhatnak a globális pozíció megtartásában. Általános megállapítás, hogy a kamerát érdemes olyan kiegészítő szenzorral társítani, amely a vizuális SLAM ismert gyenge pontjait ellensúlyozza (Cadena et al., 2016).

3. MÓDSZERTAN

A javasolt módszer alapját az ORB-SLAM3 vizuális SLAM rendszer monokuláris konfigurációja képezi. Az ORB-SLAM3 szálanként elkülöníti a képkövetési, a lokális térképezési és a hurok zárási folyamatokat, ami lehetővé teszi a valós idejű működést nagy kiterjedésű környezetekben is. A rendszer kulcsképeket és hozzájuk tartozó térképpontokat használ a környezet reprezentálására, miközben a kamera pózait egy lokálisan konzisztens pözgráfban tárolja.

A monokuláris konfiguráció sajátossága, hogy a térkép egy relatív koordináta-rendszerben épül fel, ahol a skála nem ismert és az idő előrehaladtával globális sodródás jelenhet meg. A javasolt módszer célja ezen sodródás csökkentése és a térkép abszolút, földrajzi referencia-rendszerhez történő rögzítése globális navigációs műhold-rendszer mérések felhasználásával, a vizuális SLAM belső struktúrájának megőrzése mellett.

3. 1. UTM koordináták alkalmazása pözgráf optimalizációra

A GNSS rendszer által szolgáltatott pozíciómérések az egységes térbeli referencia biztosítása érdekében az UTM vetületi koordináta-rendszerbe kerülnek átalakításra. A GNSS mérések időben szinkronizáltak a kameraképekkel, így minden kulcsképhez hozzárendelhető egy megfelelő UTM pozíció.

A vizuális SLAM által becsült kamera pozíciók és az UTM koordináták közötti kapcsolatot nem közvetlen korlátozásokkal építjük be a pözgráfba, hanem egy globális hasonlósági transzformáción keresztül írjuk le. Ez lehetővé teszi, hogy a vizuális SLAM lokális konzisztenciája sértetlen maradjon.

3. 2. UTM koordináták integrációja valós idejű vSLAM térképezés megvalósítására

A javasolt módszer célja a monokuláris vizuális SLAM rendszer által becsült pálya és térkép globális, GNSS-alapú referencia-rendszerhez történő rögzítése oly módon, hogy a valós idejű működés és a vizuális SLAM lokális geometriai konzisztenciája végig megmaradjon. Ennek érdekében a GNSS integráció nem egyetlen, teljes optimalizációs lépésben történik, hanem fokozatosan, különböző számítási igényű műveletek egymásutánjaként valósul meg. A vizuális SLAM által becsült pálya és a GNSS mérésekből származó UTM koordináták közötti kapcsolatot egy globális hasonlósági transzformáció leképezés írja le, amelyek kezdeti becslését az Umeyama módszerrel (Umeyama, 1991) határozzuk meg a kulcsképek pozíciói alapján.

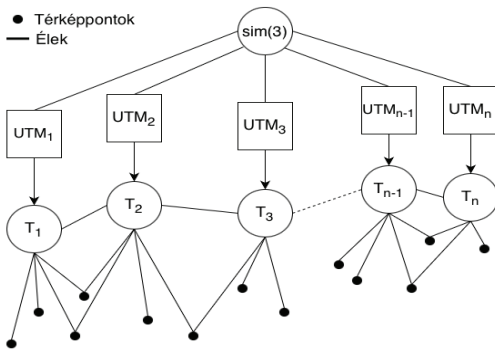
A térképezési folyamat során a GNSS információkat valós időben kizárólag egy korlátozott szabadsági fokú és paraméterű póz-igazítás formájában használjuk fel. Ebben a fázisban a hasonlósági transzformáció rögzített paraméterként szerepel, és az optimalizáció kizárólag a kulcskép-pózokra hat. A térképpontok nem vesznek részt az optimalizációban, azok a kulcsképekhez viszonyítva rögzítettnek tekinthetők, ezért ez a lépés nem tekinthető klasszikus köteget kiegészítésnek.

Legyen $T_i \in SE(3)$ az i -edik kulcskép póza, $u_i \in \mathbb{R}^3$ a hozzá tartozó GNSS mérés UTM koordinátában, valamint $S \in Sim(3)$ a rögzített globális hasonlósági transzformáció. A kulcskép UTM térbe vetített pozíciója $S \circ \text{trans}(T_i)$ alakban adódik, és az optimalizáció célja ezen pozíciók illesztése a GNSS mérésekhez a következő költségfüggvény minimalizálásával:

$$\min_{\{T_i\}_{i \in \mathcal{K}}} \sum_{i \in \mathcal{K}} \rho \left(\| u_i - S \circ \text{trans}(T_i) \|_{\Sigma_i}^2 \right),$$

ahol \mathcal{K} a GNSS-sel ellátott kulcsképek halmaza, Σ_i az adott GNSS mérés kovariancia mátrixa, és $\rho(\cdot)$ robusztus veszteségfüggvény. A számítási igény korlátozása érdekében ezt az igazítást jellemzően öt kulcsképenként futtatjuk, ami már a térképezés során is jelentősen csökkenti a vizuális odometria globális torzulásait, miközben nem veszélyezteti a valós idejű működést. Az UTM-alapú póz-igazítást minden esetben egy

lokális kötegelt igazítás követi, amely kizárólag képi megfigyelésekre támaszkodva finomítja a lokális térképpont-struktúrát, biztosítva a vSLAM belső konzisztenciáját és a stabil leíró követés fenntartását.



1. ábra: A tervezett módszer faktorgráfja UGBA esetén. A térképpontok és a GNSS alapú UTM mérések is részt vesznek az optimalizációban. A globális hasonlósági transzformációt a „sim(3)” elnevezésű csúcs jelöli (Forrás: saját szerkesztés)

A térképezési folyamat lezárását követően egy teljes, nem valós idejű UTM-alapú globális kötegelt kiegyenlítést futtatunk (1. ábra), amely a végleges térkép előállítását szolgálja. Ebben a fázisban a korábban rögzített hasonlósági transzformáció felszabadításra kerül, és a kulcsképpozók mellett a térképpontok is részt vesznek az optimalizációban (UGBA). Ez különösen indokolt, mivel az Umeyama módszerrel számított kezdeti hasonlósági transzformáció a kulcsképek kezdeti pozícióira épül, miközben az UGBA során a kulcsképpozók fokozatosan finomodnak. A globális hasonlósági transzformáció optimalizálása így lehetővé teszi, hogy a GNSS-alapú abszolút referencia jobban illeszkedjen a már finomított vizuális struktúrához. A teljes UGBA során az optimalizációs probléma a következőképpen írható fel:

$$\min_{\mathbf{S}, \{\mathbf{T}_i, \mathbf{X}_j\}} \sum_{(i,j) \in \mathcal{O}} \|\mathbf{z}_{ij} - \pi(\mathbf{T}_i, \mathbf{X}_j)\|_{\Omega_{ij}^{-1}}^2 + \sum_{i \in \mathcal{K}} \rho(\|\mathbf{u}_i - \mathbf{S} \circ \text{trans}(\mathbf{T}_i)\|_{\Sigma_i^{-1}}^2),$$

ahol \mathbf{X}_j a térképpontok, \mathbf{z}_{ij} a képi megfigyelések, $\pi(\cdot)$ a kamera vetítési modellje, míg Ω_{ij} és Σ_i rendre a vizuális és GNSS mérések információs mátrixai. Ebben a végső optimalizációs lépésben a vizuális és GNSS-alapú információk együttesen kerülnek figyelembevételre, amely globálisan konzisztens, abszolút koordináta-rendszerhez rögzített térképet eredményez. A kész térképet a becsült hasonlósági transzformációval együtt mentjük el, ami lehetővé teszi, hogy egy másik jármű saját, előzetesen kalibrált kamerájával az ORB-SLAM3 belső relokalizációs algoritmusát használva vizuálisan lokalizáljon, miközben a póz közvetlenül UTM koordináta-rendszerben is meghatározható.

4. EREDMÉNYEK

A tervezett rendszer kiértékeléséhez egy saját készítésű adatszettet használtunk, amely a ZalaZone tesztpálya közvetlen környezetében került rögzítésre az általunk konfigurált mérőjármű alkalmazásával. A mérőjármű a 2. ábrán látható. Az adatszett 20 Hz frekvenciával rögzített, 2 megapixel-es kamera képeket, valamint PTP-alapú időszinkronizációval ellátott, 100 Hz-es mintavételezésű, hozzávetőleg ± 5 cm pontosságú műholdas pozícióméréseket tartalmaz.

A kijelölt útvonalon a mérést kétszer végeztük el: az első bejárás a térképezési fázist, míg a második a lokalizáció vizsgálatát szolgálta. Ez a felosztás lehetővé tette a térkép és a lokalizáció teljesítményének egymástól független vizsgálatát.

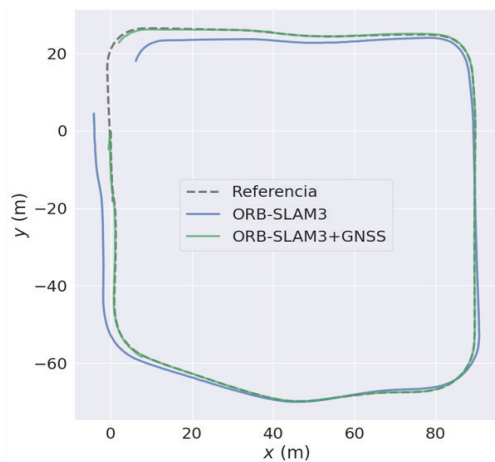


2. ábra: Felszenzorozott mérőjármű

A bemutatott módszert egyrészt valós időben, a helyszínen is vizsgáltuk, másrészt a felvett adatokat ROS2 (Macenski et al., 2022) adatszettként rögzítettük, és utólagos módon visszajátszottuk a részletes kiértékeléshez. A méréseket mind a tervezett UTM-alapú GNSS integráció alkalmazásával, mind annak mellőzésével elvégeztük, így közvetlen összehasonlításra nyílt lehetőség.

Az eredmények alapján megállapítható, hogy a folyamatos GNSS (UTM) integráció már a térképezési folyamat során is érdemben javította a pózgráf globális pontosságát. A végső, GNSS-asszisztált globális optimalizáció (UG-BA) után kapott pálya további, egyértelmű javulást mutatott. A 3. ábrán zöld színnel látható a GNSS-sel korrigált útvonal, míg késsel az alap ORB-SLAM3 által becsült pálya.

Az útvonalak pontosságát az abszolút póz hiba (APE) kiértékelési módszerrel értékeltük, amely a referencia és a becsült pálya időben illesztett pontjai közötti eltérést méri. A kiértékelés során az evo nyílt forráskódú eszközkészlet segítségével számoltuk ki az APE értékeket és vizualizáltuk a pályákat (Grupp, 2017). A 3. ábrán szemléletesen is látható, hogy a javasolt módszerrel a pálya lényegesen pontosabban követi a referencia trajektóriát. A 2. táblázat a térképezési fázis számszerű eredményeit foglalja össze. Ezek alapján megállapítható, hogy a GNSS mérések integrálásával a pálya globális deformációja jelentősen csökkenthető: az átlagos APE értéke



3. ábra: A térkép készítése során bejárt útvonal pályája GNSS használata nélkül késsel és GNSS használata esetén zölddel jelölve

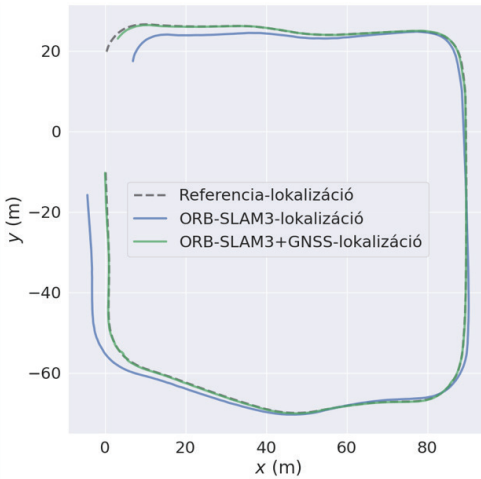
több mint harmadára csökkent az eredeti ORB-SLAM3 megoldáshoz képest. Ennek eredményeként a létrejövő vSLAM térkép tiszta lokalizációs célú újrafelhasználása is egy lényegesen pontosabb térbeli reprezentáción alapul.

Mérték[m]	ORB-SLAM3	ORB-SLAM3 + GNSS
Maximum	6.88	4.06
Átlag	2.41	0.20
Medián	1.87	0.14
Minimum	0.55	0.02
RMSE	2.90	0.35
SSE	13851.29	205.63
Szórás	1.61	0.29

2. táblázat: Az APE eredmények számszerűsítve térképkészítés során. A táblázatban jelölt hibák mértékegysége: [m].

A további kiértékelés során megvizsgáltuk, hogy a globális pozíciómérések felhasználásával készített térképen történő lokalizáció megbízhatóan működik-e, és a várt pontos lokalizációt eredményezi-e. A lokalizáció pontosságának vizsgálatához, a korábban elkészített térképre szkenelt képi és globális pozíció adatokat rögzítettünk, közel ugyanazon az úton haladva, mint a térkép készítése során. A vizsgálat során elemeztük, hogy a GNSS mérések felhasználásával készült térkép mennyire alkalmas pontos és megbízható helyzetmeghatározásra. A rögzített kamera képeket ROS2 keretrendszeren keresztül továbbítottuk a vSLAM algoritmus felé, amelyet a korábban elkészített térkép betöltésével indítottunk. A ROS2 keretrendszer alkalmazása biztosította az utólagos, de mégis valós idejű tesztelési környezetet.

A felvétel elindítása után a rendszer automatikusan relokalizálta magát, majd lokalizációs módba kapcsolott. Ebben az üzemmódban a vSLAM nem bővíti tovább a térképet, nem ad hozzá új kulcsképeket vagy térképpontokat, hanem kizárólag a meglévő térkép elemei alapján becsüli az aktuális kamera-pózt. A térképhez eltárolt hasonlósági transzformációs mátrix felhasználásával az aktuális kameraképekhez tartozó UTM koordinátákat meghatároztuk. Vagyis a vSLAM által becsült pózt eltranszformáltuk a hasonlósági transzformációs mátrix-szal. Az így kapott UTM koordinátákat összevetettük az adatszettben referenciaként rögzített képekhez tartozó UTM koordinátákkal.



4. ábra: A lokalizáció vizuális eredménye a GNSS felhasználása nélkül kékkel, és a GNSS mérések használata esetén zölddel jelölve

A lokalizáció számszerű eredményeit a 3. táblázat, míg a vizuális összehasonlítást a 4. ábra mutatja be. Az eredmények egyértelműen igazolják, hogy a GNSS mérésekkel korrigált térkép használata nemcsak a térképezés pontosságát javítja, hanem a későbbi lokalizáció során is szignifikánsan kisebb hibát eredményez.

Mérték[m]	ORB-SLAM3-lok.	ORB-SLAM3 + GNSS-lok.
Maximum	6.97	4.30
Átlag	2.45	0.255
Medián	1.83	0.15
Minimum	0.46	0.03
RMSE	2.93	0.49
SSE	16904.58	489.08
Szórás	1.61	0.41

3. táblázat: Számszerűsített eredmények a lokalizációs vizsgálat során, ahol a készített térkép alapján történt a helyzetmeghatározás

5. KONKLÚZIÓ

A bemutatott módszer alapján megállapítható, hogy a tervezett módszer képes jelentősen csökkenteni a monokuláris vSLAM rendszerekre jellemző skálahibát, ezáltal pontosabb pozícióbecslést és megbízhatóbb környezeti térképet biztosít. A globálisan skálázott és UTM koordináta-rendszerbe illesztett térkép más járművek számára is újrafelhasználható lokalizációs célra,

így lehetőség nyílik olyan kooperatív rendszerek kialakítására, amelyekben nem minden járműnek szükséges nagy pontosságú, költséges GNSS-RTK rendszerrel rendelkeznie – ez a megközelítés érdemi költségcsökkentést eredményezhet flottaszintű alkalmazásokban.

A jelen vizsgálat során elsősorban statikus környezetben, többnyire állandó fényviszonyok mellett került sor a rendszer értékelésére. A dinamikus objektumok (pl. járműforgalom) és a környezeti változások (pl. időjárás, megvilágítás) hatását a térkép megbízhatóságára jelen tanulmány nem tárgyalta. Ugyanakkor az elvégzett kísérletek azt mutatják, hogy a javasolt módszer stabil és pontos lokalizációt képes biztosítani azonos vagy hasonló körülmények között, mint amelyek során a térképezés készült – ez pedig megfelelő alapot nyújthat a jövőbeni, kiterjesztett alkalmazások számára.

KÖSZÖNETNYÍLVÁNÍTÁS

A kutatást a Kulturális és Innovációs Minisztérium támogatta a Nemzeti Kutatási, Fejlesztési és Innovációs Alap keretében, a „Nemzeti Laboratóriumok Pályázati Program” támogatási konstrukcióján keresztül, a 2022-2.1.1-NL-2022-00012 számú projekt keretében.

FELHASZNÁLT IRODALOM

- [1] Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J. J. (2016) Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics*, 32 (6), pp. 1309–1332. DOI: <https://doi.org/10.1109/TRO.2016.2624754>
- [2] Campos, C., Elvira, R., Gómez Rodríguez, J. J., Montiel, J. M. M., Tardós, J. D. (2021) ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multi-map SLAM. *IEEE Transactions on Robotics*, 37 (6), 1874–1890. DOI: <https://doi.org/10.1109/TRO.2021.3075644>
- [3] Chen, X., Miao, Y., Wang, J., Li, Y. (2018) GVORB-SLAM: A GPS-aided visual SLAM system. *ISPRS International Journal of Geo-Information*, 7 (11), 425. DOI: <https://doi.org/10.36227/techrxiv.23898246>

- [4] Geiger, A., Lenz, P., Stiller, C., Urtasun, R. (2012) Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research*, 32 (11), pp. 1231–1237. <https://doi.org/10.1177/0278364913491297>
- [5] Grupp, M. (2017) evo: Python package for the evaluation of odometry and SLAM [Computer software]. GitHub repository. URL: <https://github.com/MichaelGrupp/evo>
- [6] Kiss, B., Illés, G., Tóth, C. K. (2019) GPS-SLAM: Global localization using visual SLAM and GPS measurements. *Sensors*, 19 (12), 2740.
- [7] Leutenegger, S., Lynen, S., Bosse, M., Siegwart, R., Furgale, P. (2015) Keyframe-based visual-inertial odometry using nonlinear optimization. *International Journal of Robotics Research*, 34 (3), pp. 314–334. DOI: <https://doi.org/10.1177/0278364914554813>
- [8] Macenski, S., Foote, T., Gerkey, B., Lalanette, F., Woodall, W. (2022) Robot Operating System 2: Design, architecture, and uses in robotics. *Science Robotics*, 7 (66), eabm6074. DOI: <https://doi.org/10.1126/scirobotics.abm6074>
- [9] Mur-Artal, R., Montiel, J. M. M., Tardós, J. D. (2015) ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31 (5), pp. 1147–1163. DOI: <https://doi.org/10.1109/TRO.2015.2463671>
- [10] Mur-Artal, R., Tardós, J. D. (2017). ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Transactions on Robotics*, 33 (5), pp. 1255–1262. DOI: <https://doi.org/10.1109/TRO.2017.2705103>
- [11] Qin, T., Li, P., Shen, S. (2018) VINS-Mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34 (4), pp. 1004–1020. DOI: <https://doi.org/10.1109/TRO.2018.2853729>
- [12] Scaramuzza, D., Fraundorfer, F. (2011) Visual odometry [Tutorial]. *IEEE Robotics & Automation Magazine*, 18 (4), 80–92. DOI: <https://doi.org/10.1109/MRA.2011.943233>
- [13] Wang, Y., Zhang, J., Zhao, H. (2022) Map reuse and cooperative localization for multi-agent SLAM systems. *Robotics and Autonomous Systems*, 148, 103922.
- [14] Zhang, J., Singh, S. (2014) LOAM: Lidar odometry and mapping in real-time. In *Proceedings of Robotics: Science and Systems (RSS)*. DOI: <https://doi.org/10.15607/RSS.2014.X.007>
- [15] Zhang, Z., Qin, T., Shen, S. (2019) LD-SLAM: A robust and efficient monocular SLAM system based on line and point features. *IEEE Robotics and Automation Letters*, 4 (2), pp. 1378–1385.



Cooperative visual SLAM system for vehicle based environmental perception

Keywords: localization, GNSS-integration, cooperative vSLAM, autonomous vehicle

For autonomous vehicles, accurate assessment of the environment and reliable localization are crucial. To do this, we present a cooperative visual SLAM system, where one vehicle creates a map with a monocular camera and calibrates this map to a global scale using GNSS reference coordinates. The second vehicle visually localizes itself using the created map interpreted in the global coordinate system. We used a self-generated dataset to evaluate the presented method. Based on the results, it can be concluded that the method increases the accuracy of localization and allows the reuse of maps between vehicles.