

Peremhálózat alapú VSLAM keretrendszer erőforrás-korlátozott platformokhoz

An Edge Assisted VSLAM Framework for Resource Constraint Platforms

Prof. Dr. VÁSÁRHELYI József¹, SALIH M. Omar²

Miskolci Egyetem, Gépészmérnöki és Informatikai Kar, Automatizálási és Infokommunikációs Intézet,
3515 Miskolc, Egyetemváros
{josef.vasarhelyi,omar.salih}@uni-miskolc.hu}

Abstract – Localizarea și cartografierea simultană vizuală (VSLAM) joacă un rol fundamental în capacitatea roboților mobili de a localiza și de a construi hărți ale mediului înconjurător. Cu toate acestea, cerințele computaționale intensive ale VSLAM limitează implementarea sa pe platforme cu resurse limitate. Această lucrare prezintă un cadru VSLAM asistat de edge computing, care împarte sarcina de calcul între roboții mobili și un dispozitiv edge centralizat.

Abstract -- Visual Simultaneous Localization and Mapping (VSLAM) play a fundamental role in enabling mobile robots to localize and construct maps of their surrounding environments. However, the intensive computational demands of VSLAM restrict its deployment on resource-constrained platforms. This paper introduces an edge-assisted VSLAM framework that partitions the computational workload between mobile robots and a centralized edge device.

Keywords: localization and mapping, data compression, edge computing, resource constraint devices.

Kivonat – A vizuális egyidejű helymeghatározás és térkép készítés (VSLAM) alapvető szerepet játszik abban, hogy a mobil robotok képesek legyenek meghatározni helyzetüket és térképet készítsenek környezetükről. Azonban a VSLAM számításigényes jellege korlátozza annak alkalmazását erőforrás-korlátozott platformokon. Jelen cikk egy peremhálózat alapú VSLAM keretrendszert mutat be, amely a számítási terhelést megosztja a mobil robot és egy központi edge - eszköz között.

Kulcsszavak: helymeghatározás, térkép készítés, peremhálózati számítás, adattömörítés, erőforrás-korlátozott eszközök.

1. BEVEZETÉS

A VSLAM fő célja, hogy a mobil robotok önállóan fedezzék fel és járják be az ismeretlen környezetet, miközben az érzékelők által rögzített adatok segítségével készítsenek térképet és határozzák meg saját helyzetüket. A robotok autonóm navigációja és interakciója emberi környezetben csökkenti a munkaerőigényt és növeli az általános termelékenységet [1]. A mobil robotoktól gyűjtött információk széles körű alkalmazásokhoz használhatók, például felfedezési stratégiához, útvonaltervezéshez, térképek összeállításához és egyéb magasabb szintű döntésekhez. A VSLAM rendszerek azonban jellemzően nagy számítási kapacitást igényelnek. A sávszélesség korlátai és a valós idejű követelmények miatt a teljes VSLAM folyamat mobil eszközökre történő áthelyezése nem praktikus a csökkentet erőforrású eszközök esetében. Ezért a robot tömegének csökkentése, az energia megtakarítás és a mobil eszközök méretének minimalizálása érdekében a legújabb kutatások azt vizsgálják, hogyan lehetne a számításigényes VSLAM feldolgozási feladatokat edge cloud platformokra áthelyezni [2]. Ebben az esetben az összetettebb feladatok számítása kerül áthelyezésre, míg az alapvető funkciók továbbra is a mobil eszközön maradnak.

A kutatások különféle tömörítési stratégiákat vizsgáltak annak érdekében, hogy javítsák az adatátvitel hatékonyságát és kezeljék a sávszélesség korlátait az edge-alapú VSLAM megvalósításokban. Az egyik megközelítés JPEG-alapú algoritmusokat alkalmazott a pontfelhő-adatok tömörítésére, a megosztott környezeti felvételeket panorámaképekké alakította [3].

Martins és munkatársai a tömörítési módszerek összehasonlító elemzésével optimalizálták az adatmegosztást többrobotos rendszerekben. A [4] cikk az elfoglaltsági ráctérképek tömörítésére adott választ az együttműködő SLAM alkalmazásokban. Eredményeik szerint a szabványos tömörítési algoritmusok gyakorlati megoldást jelentenek a sávszélesség csökkentésére [4].

Zhang és Deng a többrobotos SLAM kommunikáció által felvetett problémákat vizsgálták, és veszteségmentes kódolási stratégiákat alkalmaztak. Az adatátvitel előtt módosított Hoffmann - kódolást alkalmaztak futáshossz – plusz kódolással együtt a teljes térkép tömörítése érdekében, miközben kompakt CNN-eket integráltak a jellemzők kinyerésére [5].

Liu és munkatársai kizárólag a kulcsképkockák jellemző tulajdonságainak tömörítésére összpontosítottak egy központosított ORB-SLAM3 keretrendszerben [6].

Zhong és munkatársai kvantáláson alapuló redundáns kulcsképkocka - eliminációs módszert fejlesztettek ki, hogy csökkentsék a túlzott adatmennyiségből eredő feldolgozási késleltetéseket, és javítsák a valós idejű rendszer teljesítményét [7].

Li és munkatársai a CORB-SLAM keretrendszer egyrobotos ORB-SLAM2 képességeit terjesztette ki többrobotos környezetekre, memóriakezelő modulok integrálásával. Az alkalmazott módszer összehangolja a lokális térkép készítést és a szerverkommunikációs feladatokat[8]. Az edge SLAM [9] az edge computing alkalmazásával próbálta megoldani a mobil eszközök erőforrás-korlátait csökkentve a kommunikációs késleltetéseket. A képfeldolgozás három párhuzamos kommunikációs csatornát alkalmazott a jellemző adatok, geometriai adatok, kulcsképkocka-jelöltek, valamint lokális térképfriessítések továbbítására.

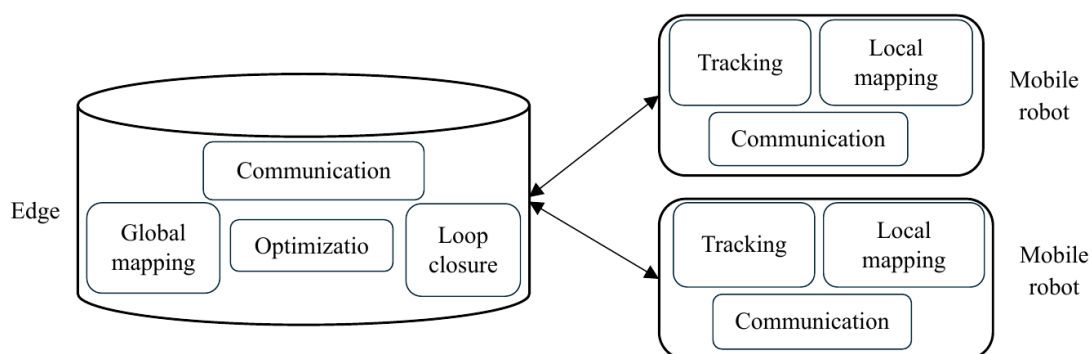
Bár számos kutatás foglalkozott vizuális elemek – például jellemzők és térképpontok – veszteségmentes tömörítésével, a veszteséges tömörítési módszerek potenciálisan nagyobb hatékonyságot kínálnak a VSLAM alkalmazások számára.

A cikkben az általánosított, ORB-SLAM alapú, központosított edge-alapú többrobotos architektúra kerül bemutatásra. Esettanulmányként két mobil robotot használunk. A fő VSLAM feladatokat megosztottuk a mobil robotok és az szerver eszköz között. Bemutatunk egy kommunikációs modult, amely javítja a tömörítési algoritmust. Elemezzük a vizuális adatok veszteséges tömörítésének hatásait a V-SLAM működésében, különös tekintettel a robusztusságra és a pályapontosságra.

2. MÓDSZERTAN

2.1. A rendszer felépítése

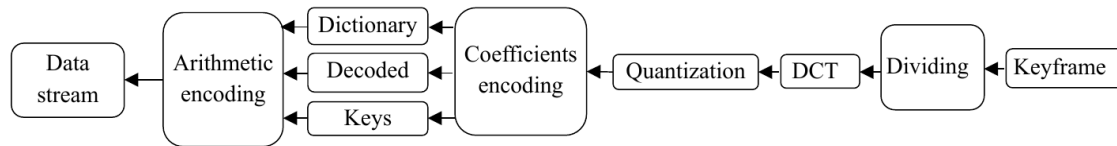
A központosított, edge-alapú VSLAM keretrendszer képes a globális térkép rekonstrukciójára. A rendszerben minden robot tartalmaz követési, lokális térkép készítési és kommunikációs modult. A robotok a modulok segítségével feldolgozzák és továbbítják az adatokat az edge eszköz felé. Az edge számítógép végzi a térkép optimalizálását, egyesíti a két robot által készített térképet egy globális térképpé. Az 1. ábra szemlélteti a kidolgozott többrobotos ORB-SLAM architektúrát, amely két robotot alkalmaz, de könnyen bővíthető továbbiakkal. Az ORB-SLAM részletes magyarázatát az olvasó a [10] cikkben találja meg.



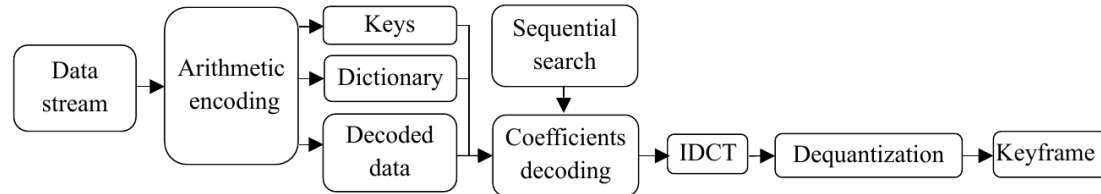
1. ábra. A javasolt rendszer tömbvázlata

2.2. Adattovábbítási rendszer

A tervezett kommunikációs rendszer két különálló komponensből áll: egy kódoló egységből és egy dekódoló egységből. A kódoló egységet a mobil robot vezérli. Az modul feldolgozza a követő egység által generált kulcsképkocka-jelölteket. A kódoló tervezési szempontjai közé tartozik az egyszerűség, a minimális



a. Kommunikációs egység – robot oldal



b. Kommunikációs egység – szerver oldal

2. ábra. A javasolt rendszer tömbvázlata

feldolgozási késleltetés a valós idejű működés fenntartása, valamint az alacsony erőforrásigény. A 2. ábra bemutatja a javasolt kommunikációs keretrendszer alapvető tömbvázlatát.

A kódolási folyamatban a bemeneti kulcsképkockákat blokkokra osztjuk fel, majd diszkrét koszinusz transzformációt (DCT) alkalmazunk, a frekvenciatartományba történő átalakítás érdekében. A frekvenciakomponensek alapján egy módosított kvantálási séma kerül alkalmazásra. Az optimális tömörítési teljesítmény elérése érdekében a rendszer egy veszteségmentes együttható-kódolási (Coefficient Encoding – CE) algoritmust használ. A CE algoritmus három véletlenszerű kulcsot alkalmaz, hogy a DCT együttható-hármasokat kódolja szorzatok összegén alapuló műveletekkel. A valószínűségi eloszlásokat tároljuk a későbbi dekódolási folyamatokhoz.

Az szerver oldali dekódolás visszafordítja ezt a műveletsort. Aritmetikai dekódolással megkapjuk a kódolt adatokat, a kulcsokat és a valószínűségi táblázatokat. Szekvenciális keresési eljárással rekonstruáljuk a DCT együtthatókat a tárolt kulcsok és szótárak segítségével. Bár ez a folyamat számításigényes, az edge computing képességeit kihasználva hatékonyan megvalósítható. A visszanyert együtthatókat dekvantálásnak és inverz DCT transzformációnak vetjük alá, hogy visszaállítsuk a térbeli reprezentációt.

3. EREDMÉNYEK

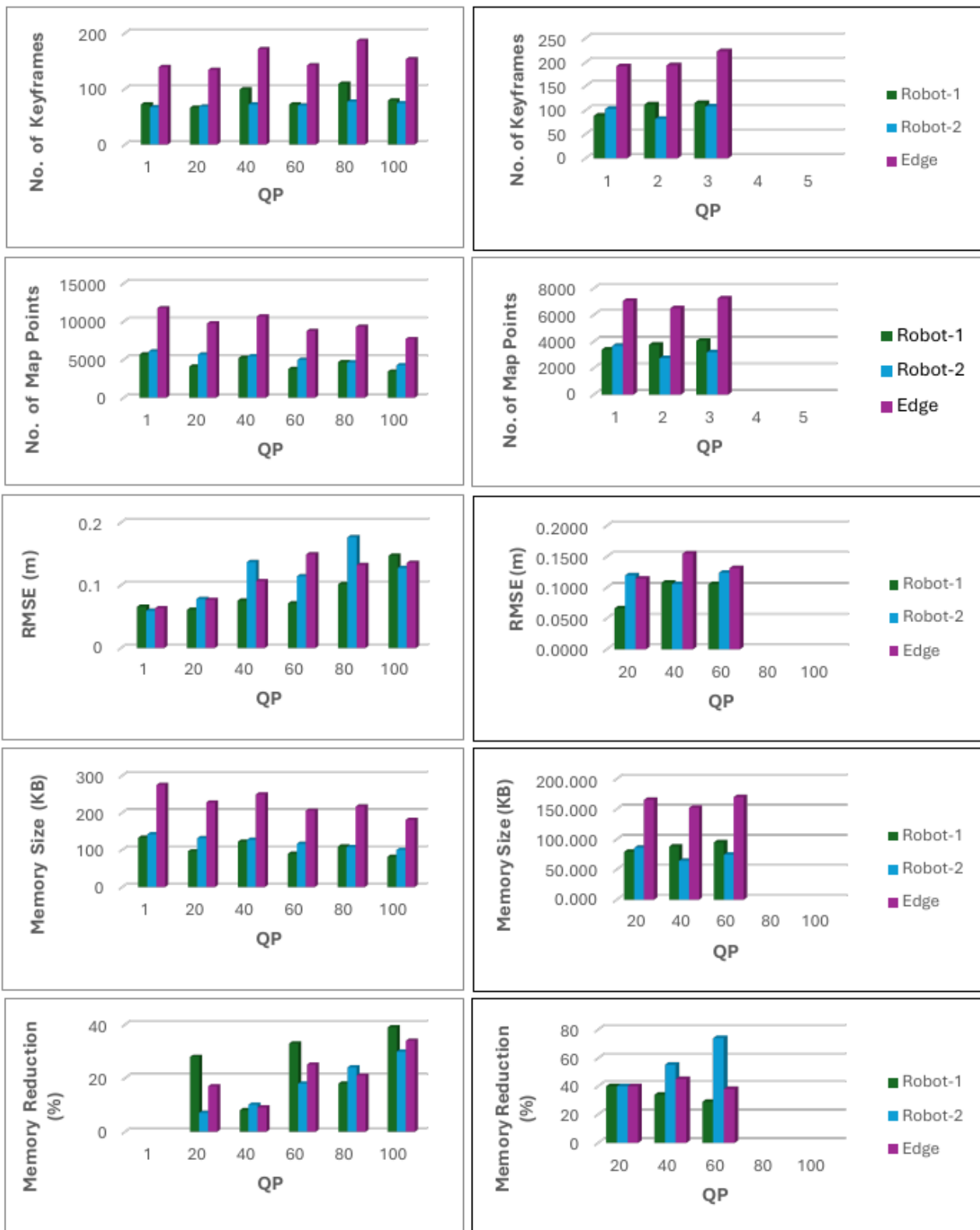
A javasolt architektúra teljesítményének értékeléséhez a TUM adatbázis „freiburg3 long office household” [11] nevű szekvenciáját használtuk. Az adatállomány 2 585 képkockát tartalmaz 640×480-as 8 bites formátumban, amely egy 21,455 m hosszú, 87,10 s időtartamú valós pályát ábrázol egy 5,12 m × 4,89 m × 0,54 m méretű térben. A képkockák elosztása a robotok között megfelelő átfedéssel történik, hogy lehetővé tegye a globális térkép integrációját és a hurokzárások felismerését.

A rendszer kiértékelésére kétfázisú módszertant alkalmaztunk. Először a többrobotos architektúrát szimuláltuk az eredeti felbontású képkockákkal, hogy felmérjük a rendszer teljesítményét és a globális térkép egyesítési képességeit. Ezt követően a kommunikációs keretrendszert vizsgáltuk különböző kvantálási paraméterek (QP) alkalmazásával, hogy meghatározzuk a kommunikációs modul hatékonyságát.

A pályapontosság kiértékeléséhez a négyzetes középhiba (RMSE) szolgál fő mérőszámként, amely az egyes szimulációs lépések során becslt és valós pályák közötti abszolút eltérést méri. A kifejlesztett rendszer teljesítményét összehasonlítottuk egy JPEG-alapú referencia konfigurációval, amely a JPEG-et használja szabványos tömörítési eljárásaként.

A JPEG-alapú rendszer megbízható teljesítményt mutat $Q \leq 60$ paramétertartományban; azonban ennek a küszöbértéknek a túllépése a globális térkép egyesítésének hibájához vezet, amit a pályabecslés romlása okoz a JPEG tömörítési hibák miatt. Ezzel szemben a javasolt keretrendszer jobb teljesítményt ér el a JPEG-alapú megvalósításhoz képest mind a tömörítési hatékonyság, mind a rekonstruált képkockák minősége tekintetében, egészen $Q=100$ kvantálási értékig. Ezen érték fölött azonban a javasolt rendszer is hibát mutat a pályabecslés és a globális térkép készítés terén.

A javasolt rendszer kódolási ideje lassabb a JPEG-nél a többlet kódolási lépések miatt, ugyanakkor ez a számítási többlet nagyobb tömörítés hatékonysággal közelít a referencia teljesítményhez.



3. ábra. Eredmények összehasonlítása – bal: javasolt rendszer, jobb: JPEG alapú rendszer

A 3. ábra bemutatja az eredményeket összehasonlítva a JPEG alapú rendszerrel. Tartalmazza a pályabecslés pontosságát. Szemlélteti a rendszer hatékonyságát alacsony és magas kvantálási szintek mellett. A kísérleti eredmények azt mutatják, hogy a javasolt keretrendszer robusztus V-SLAM működést tesz lehetővé mindkét robotplatformon, és nagy pályapontosságot ér el az RMSE elemzés alapján. Emellett az edge-alapú feldolgozás sikeresen végzi el a globális térkép egyesítését a két robot pozícióbecslései alapján, amelyek még magas kvantálási értékek mellett is megbízhatóak és megfelelő működést biztosítanak, konvergáló RMSE teljesítményt eredményezve.

IRODALMI HIVATKOZÁSOK

- [1] Macario Barros, A., Michel, M., Moline, Y., Corre, G., & Carrel, F. (2022). A Comprehensive Survey of Visual SLAM Algorithms. *Robotics*, 11(1), 24.
- [2] Hofer, Johannes, Peter Sossalla, Christian L. Vielhaus, Justus Rischke, Martin Reisslein, and Frank HP Fitzek. "Comparison of Analyze-Then-Compress Methods in Edge-Assisted Visual SLAM." *IEEE Access* (2023).
- [3] L. Zheng, K. Xu, J. Jiang, M. Wei, B. Zhou, and H. Cheng, "Real-time efficient environment compression and sharing for multi-robot cooperative systems," *IEEE Trans. Intell. Veh.*, vol. 10, no. 1, pp. 1–14, 2024.
- [4] G. S. Martins, D. Portugal, and R. P. Rocha, "A comparison of general-purpose FOSS compression techniques for efficient communication in cooperative multi-robot tasks," in *Proc. 11th Int. Conf. Informat. Control, Autom. Robot. (ICINCO)*, vol. 2, 2014, pp. 136–147.
- [5] L. Zhang and J. Deng, "Deep compressed communication and application in multi-robot 2D-Lidar SLAM: An intelligent Huffman algorithm," *Sensors*, vol. 24, no. 10, p. 3154, 2024.
- [6] X. Liu, S. Wen, J. Zhao, T. Z. Qiu, and H. Zhang, "Edgeassisted multi-robot visual-inertial SLAM with efficient communication," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, no. 4, pp. 2186–2198, 2024.
- [7] R. Zhong, Y. Bai, A. Wang, Z. Hu, and H. Fang, "Centralized multi-robot visual SLAM and scene reconstruction," in *Proc. 43rd Chinese Control Conf. (CCC)*, 2024, pp. 3748–3754.
- [8] F. Li, S. Yang, X. Yi, and X. Yang, "Corb-slam: A collaborative visual slam system for multiple robots," in *Proc. 13th Int. Conf. Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, 2017, pp. 480–490.
- [9] A. J. Ben Ali, M. Kouroshli, S. Semenova, Z. S. Hashemifar, S. Y. Ko, and K. Dantu, "Edge-SLAM: Edge-assisted visual simultaneous localization and mapping," *ACM Trans. Embedded Comput. Syst.*, vol. 22, no. 1, pp. 1–31, 2022.
- [10] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *Proc. Int. Conf. Intelligent Robot Systems (IROS)*, 2012, pp. 573–580.
- [11] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB_SLAM: A versatile and accurate monocular SLAM system," *IEEE Trans. Robot.*, vol. 31, no. 5, pp. 1147–1163, 2015.