




Az útfelület normálvektorának predikciója és gyakorlati alkalmazásai

Markó Norbert^{1,2*}, Ballagi Áron², Szirányi Tamás^{1,3}

¹HUN-REN SZTAKI, Gépi Érzékelés Kutatólaboratórium, Budapest, Magyarország

²Széchenyi István Egyetem, Járműipari Kutatóközpont, Győr, Magyarország

³Budapesti Műszaki és Gazdaságtudományi Egyetem, Budapest, Magyarország

*Levelező szerző, e-mail: marko.norbert@ga.sze.hu

Beérkezett: 2025. augusztus 10.; elfogadva: 2025. október 21.

Összefoglalás

Tanulmányunk összefoglalja korábbi cikkeink eredményeit az útfelület normáljának monokuláris kamerából történő becsléséről. A vizsgált probléma a lejtők, emelkedők és hirtelen átmenetek okozta torzulások kezelése. Módszerünk képpáralapú homográfián alapul, amelyet a késői IMU-fúzióval viláskoordinátába forgatunk, illetve időben SLERP-szűrővel stabilizálunk. Az algoritmust a PandaSet, KITTI és saját GradeSet adatkészleteken értékeltük ki, ahol újralibrálás nélkül, következetesen alacsony normál- és pitch-hiba-értékeket kaptunk a referenciához és az előző state-of-the-art algoritmushoz képest. A kutatásaink során kidolgozott megközelítés robusztus, jól általánosítható alapot ad útfelület normálisának kameraalapú becsléséhez, ami számos további gyakorlati alkalmazásban jól hasznosítható.

Kulcsszavak: útfelület-normál, homográfia, IMU-fúzió, SLERP

Prediction of road surface normal vector and its practical applications

Norbert Markó^{1,2*}, Áron Ballagi², Tamás Szirányi^{1,3}

¹HUN-REN SZTAKI, Machine Perception Research Laboratory, Budapest, Hungary

²Széchenyi István University, Vehicle Industry Research Center, Győr, Hungary

³Budapest University of Technology and Economics, Budapest, Hungary

Summary

This paper summarizes our previous results on estimating the road surface normal from a monocular camera. We address distortions caused by slopes, grades, and abrupt transitions where projection errors grow quickly. A one degree tilt error at 50 m implies almost one meter vertical discrepancy, which can mislead free-space inference, 3D deprojection and obstacle reasoning. By explicitly recovering the road plane orientation we provide a strong geometric prior for camera-first perception. The pipeline consumes two consecutive frames, enhances road texture with contrast-limited adaptive histogram equalization, and computes dense correspondences with a lightweight transformer matcher (EfficientLoFTR). Robust homography is then estimated with MAGSAC over a road-only region of interest set near a reaction distance of roughly six meters. The homography is normalized with the intrinsic calibration to isolate geometry and improve numerical stability, then decomposed into rotation, translation and plane normal. Because matches are constrained to the road, the recovered normal describes the ground plane in the camera frame. We rotate this normal into a physically meaningful world frame through late fusion with IMU-based odometry, and stabilize the time series with spherical linear interpolation across unit quaternions. SLERP delivers smooth, unit-length updates without the parameter sensitivity and matrix inversions of Kalman-style filters, while remaining efficient for embedded deployment. We evaluate on PandaSet, KITTI, and our own GradeSet, a dataset that targets dynamic grade changes and steep ramps. A single set of parameters is used across all data, yet the method maintains consistently low errors. Averaged over challenging segments we measure a normal error of 1.18 degrees and a pitch error of 0.75 degrees, compared with 3.15 and 1.72 degrees for a strong reference. The approach handles low-texture asphalt, lane-marking scarcity and rapid grade transitions, and it does not require dataset-specific recalibration. Beyond accuracy, the explicit surface orientation enables precise inverse perspective mapping, strengthening free-space detection from a single moving camera. Practical payoffs include adaptive speed control on heavy ground

vehicles, which benefits from a stable estimate of slope and pitch for better comfort, energy use and safety. The predicted normal also supports continuous compensation of camera extrinsics relative to the road by combining an initial vanishing-point calibration with our per-frame normal, reducing long-term drift and improving consistency in camera-only stacks. The result is a robust, generalizable and resource-efficient basis for road-surface normal estimation. We also outline a forward path toward piecewise-plane modeling and extension to unstructured off-road terrain, where richer features and surface models can carry the same late-fusion and temporal-smoothing principles to more varied geometry while preserving the efficiency advantages of cameras.

Keywords: road surface normal, homography, IMU-fusion, SLERP

Bevezetés

A közúti, épített útfelületek normálisának meghatározása alapvető szerepet játszik abban, hogy a gépilátás-alapú környezetérzékelési és -észlelési rendszerek pontos és megbízható információt szolgáltatassanak a jármű környezetéről. A legtöbb modern észlelési architektúra nem veszi figyelembe explicit módon a burkolat síkjának orientációját, vagy ha igen, az is csak implicit módon kerül számításba. Ez azonban problémát jelenthet, mert egyfokos dőlésszöghiba ötvenméteres távolságban közel egy méter vertikális eltérést eredményezhet. Könnyen belátható, hogy ekkora eltérés komolyan félrevezetheti a környezetérzékelési és -észlelési algoritmusokat, aminek következtében az akadálydetektáló vagy vezethető útfelületet detektáló algoritmusok háromdimenziós térbe való deprojekciója komoly hibákat tartalmazhat. Habár bizonyos körülmények között az útfelület orientációja elhanyagolható, több algoritmus számára is nélkülözhetetlen plusz információt nyújt. Ez különösképpen igaz abban az esetben, ha környezetérzékelési rendszerünk teljes mértékben kamerákra épül, ahol az útburkolat síkja alapvető 3D-adatként szolgálhat. A burkolat orientációjának ismerete lehetővé teszi számos vizuális algoritmus működését.

A burkolatsík normálvektorának előrejelzése a monokuláris képekből végzett háromdimenziós jelenetrekonstrukció egyik legfontosabb építőköve (Hoiem–Efras–Hebert 2007), mivel ennek hiányában az objektumok térbeli helyzete, illetve a szabad haladási tér megbízható meghatározásának pontossága drasztikusan csökken. A burkolat orientációjának figyelembevétele lehetővé teszi az inverz perspektív leképezés pontos kivitelezését (Jeong–Kim 2016; Reiher–Lampe–Eckstein 2020; Philion–Fidler 2020; Marko–Sziranyi–Ballagi 2023), valamint a síkbeli parallaxis becslését (Irani–Anandan 1996), amelyek együttesen alapvető szerepet játszanak a szabad tér megbízható detektálásában (Pfeiffer–Franke 2010). Ezen túlmenően az útfelület orientációjának integrált felhasználása a magasabb szintű érzékelési modulok – például a kamera pozíciójának becslése (Koščeká–Zhang 2005), a háromdimenziós objektumdetektálás (Chen et al. 2015) és a szemantikus szegmentálás (Alvarez et al. 2012) – pontosságát és megbízhatóságát is növeli.

Jelen munka a monokuláris kamera és az inerciális mérőegység (IMU) késői fúziójának egyszerűségére építve

törekszik a burkolatsík normáljának nagy pontosságú visszaállítására, különös hangsúlyt fektetve arra, hogy módszerünk megbízhatóan teljesítsen változó lejtők és emelkedők esetén is.

Az útfelület térbeli orientációjának pontos meghatározása a fenti algoritmusok pontosságát és megbízhatóságát egyaránt növeli, amelyek összességében mind hozzájárulnak, hogy a gépjárművek és az autonóm járműrendszerek megbízhatóbb, pontosabb környezetérzékelést valósíthassanak meg. Ennek ellenére pontos meghatározása egyetlen, mozgásban lévő kamerával még mindig kihívást jelent, különösen, ha a jármű emelkedőn vagy lejtőkön halad keresztül, mivel a változó dőlésszögek jelentősen módosítják a kép és a valós háromdimenziós geometria megfeleltetését. Kutatásunk során erre a kihívásra próbálunk választ adni. A továbbiakban ismertetjük a jelenleg használt módszert, az elért eredményeket, illetve behatóan taglaljuk ezek gyakorlati hasznosíthatóságát.

Módszer

Célunk, hogy kizárólag egyetlen kamerából és az IMU-alapú pozícióbecslésből származó adatok felhasználásával kinyerjük az útfelület háromdimenziós orientációját, amelyet egységnyi hosszúságú normálvektor formájában fejezünk ki. Módszerünk lényege, hogy minden időpillanatban feltételezzük a globális koordináta-rendszerben elérhető kamerapózt, beleértve a kamera pontos pozícióját és orientációját, amelyet a korábban említett pozícióbecslő algoritmus biztosít. Ezt a kamerapózt késői (late-stage) szenzorfüzió révén egyesítjük az útfelület normálvektorára vonatkozó predikcióval, hogy nagy pontossággal meghatározhassuk az előttünk lévő útfelület dőlésszögét. A teljes algoritmus az 1. ábrán látható, amely alapján levezetjük a továbbiakat.

Az algoritmus fő alkotóeleme a monokuláris kamera pipeline, amely két egymást követő képkockát használ bemenetként: az aktuális i_2 és az előző i_1 képkockát. Az előfeldolgozás során adaptív hisztogram-kiegyenlítést alkalmazunk a kontraszt fokozására és az útburkolaton megjelenő finom textúrák kiemelésére (Pizer et al. 1987). A feldolgozott képpár ezután a homográfia-számítási modulba kerül. Itt egy könnyűsúlyú, transzformer-alapú matcher algoritmust használunk, amely az Efficient-LoFTR nevet kapta (Wang et al. 2024), és az eredeti LoFTR modell (Sun et al. 2021) egy jobb teljesítményű

változata. A matcher algoritmus erőssége, hogy gyengébb jellemzőkkel rendelkező útfelületen is képes sűrű megfelelések számítására – ellentétben az ORB- vagy SIFT-alapú módszerekkel –, így növelve az algoritmus robusztusságát. Mindemellett az alacsony számítási igénynek köszönhetően a teljes pipeline hatékony marad. Az említett matcherből kapott eredmények a homográfia-modul bemenetére kerülnek. A homográfia-bebecsléshez a MAGSAC algoritmust választottuk (Barath–Matas–Noskova 2019). Ez egy robusztus modellbecslési eljárás, amely a zajskála-tartományok mentén végez marginalizációt – szemben a RANSAC-kal, amely előre meghatározott küszöbértékek alapján számítja a transzformációt. A lehetséges zajskálák feletti integrálással simább, statisztikailag megalapozottabb becsléseket kapunk.

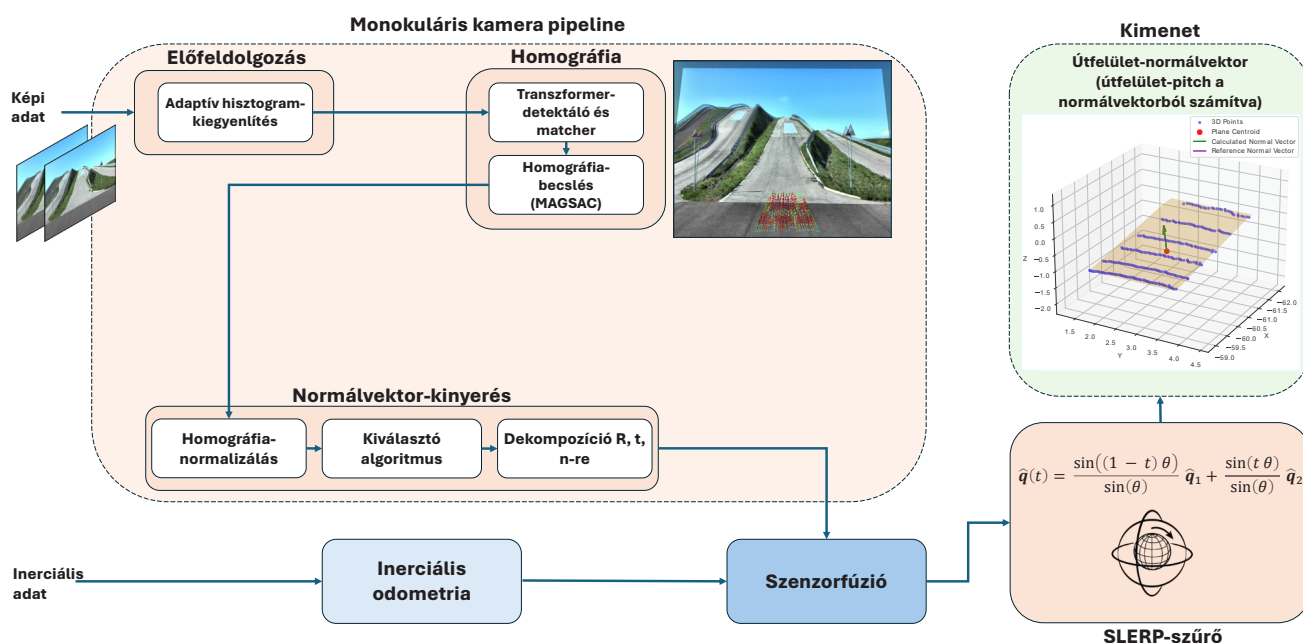
Magát a homográfia-transzformáció számítását az előttünk elterülő útfelületre korlátozzuk, hogy minimalizáljuk a kép nem úttal kapcsolatos tartományainak befolyását. Az előre meghatározott region-of-interest (ROI) a szükséges távolság alapján kerül kijelölésre. A távolság meghatározásakor a várható reakció-távolságot vettük figyelembe az adott sebesség mellett. Az esetek nagy százalékában ez nagyjából 6 méter távolságra van a járműtől.

A homográfia-számítás eredményeként kapott stabil mátrixot a normálvektor-kinyerő modulnak továbbítjuk, ahol első körben normalizálásnak vetjük alá az eredménymátrixot, amivel a kamera belső (intrinsic) paramétereinek hatását szűrjük ki, növeljük a numerikus stabilitást, és izoláljuk a transzformáció geometriai információját. Ez úgy valósul meg, hogy a homográfia-mátrixot először megszorozzuk a belső kalibrációs mátrix inverzé-

vel, majd az eredeti kalibrációs mátrixszal utólagosan szorozzuk. A normalizált homográfia-mátrix felbontható rotációs, translációs és sík-normális komponensekre. Mivel az egyezések a matcher algoritmusból az ROI-n keresztül az útburkolat síkjára korlátozódnak, a kiszámított normálvektor közvetlenül a talaj sík orientációját adja meg a kamera koordináta-rendszerében.

A késői szenzorfüzió az odometria által biztosított IMU-adatokkal elengedhetetlen ahhoz, hogy a kiszámított normálvektort fizikai értelemben konzisztens koordináta-rendszerbe forgassuk, és ezáltal növeljük algoritmusunk robusztusságát. Bár a homográfia-alapú pipeline megbízhatóan állítja elő az útfelület síkjának normálvektorát a kamera koordináta-rendszerében, ez önmagában nem képes kezelni a driftjelenségeket vagy a skála-ambiguitást. Az IMU-adatokkal történő fúzió a homográfia-bebecslés után a világkoordináta-rendszerbe transzformálja a predikciókat, így javítva az eredmények értelmezhetőségét, és megkönnyítve validálásukat.

Az időbeli konzisztencia biztosítása érdekében szűrésre is szükség volt. Ennek folyamán egy Spherical Linear Interpolation (SLERP) alapú szűrőt integráltunk a pipeline-ba (Shoemaker 1985). Ez az interpoláló algoritmus numerikusan stabil simítást nyújt az útfelület egyes normálvektorai között, alacsony számítási komplexitással, ennek nyomán beágyazott platformokon is alkalmas lehet valós idejű szűrésre. Amellett, hogy megbízható szűrést biztosít, kiküszöböli a Kálmán-szűrő-alapú módszereknél gyakran előforduló paraméter-érzékenységeket. Az előző talaj sík normálvektorát minden egyes új képkocka esetén egységkvaternióként reprezentáljuk, és ezt a kvaterniót egy rögzített interpolációs paraméter segítségével finoman elforgatjuk az újonnan megfigyelt nor-



1. ábra | A teljes algoritmus folyamatábrája. A kimeneten látható a felület normálvektora 3D-ben az adott képkocka számított területére.
 Forrás: Markó et al. 2026

málvektor-előrejelzésnek megfelelő kvaternió felé. Ez tulajdonképpen a négydimenziós egység-hiperszférán történő, a legrövidebb úton végrehajtott interpolációt valósítja meg. Mivel az interpoláció közvetlenül az egységkvaternió-szférán történik, az eredményül kapott vektorok egység-hosszúak maradnak. A módszer intuitív egyensúlyt tesz lehetővé az érzékenység és a simaság között, vagyis szabályozza, hogy az aktuális becslés mennyiben közelítse a legújabb megfigyelést a gömbi úton haladva. Ezzel szemben a Kálmán-szűrő-alapú megközelítésekhez a folyamat- és mérési zajkovariancia-mátrixok alapos hangolása szükséges. A szűrő kimenete az így finomított útfelület normálvektora, amelyből a vonatkozó pitch-szöveget is számíthatjuk. Mivel a pitch egy származtatott érték, így magában nem hordoz plusz értéket, viszont lehetővé teszi a folytonos ábrázolást és a jobb időbeli értelmezhetőséget.

Eredmények

Az alábbiakban összefoglalunk néhány, a kutatás során született eredményt, amelyek korábban tudományos konferenciákon és szakfolyóiratokban kerültek bemutatásra. Mivel jelen munka nem önálló, új kutatási projektet ismertet, inkább áttekintést kíván nyújtani a korábbi vizsgálatok alapján kifejlesztett pipeline teljesítményéről. Az eredmények bemutatása nem a kísérleti részletekre, hanem a rendszer működésének átfogó értékelésére helyezi a hangsúlyt. Ennek megfelelően cikkünk elsősorban a gyakorlati alkalmazhatóság és a társadalmi hasznosíthatóság szempontjaira fókuszál – ez a diszkusszióban kap helyet, ahol kielemezzük, hogy a fejlesztett módszer miként járulhat hozzá a valós környezetben jelentkező problémák megoldásához és a technológia szélesebb körű adaptációjához.

Kutatásaink során – ahogyan azt a bevezetőben is kiemeltük – a fő hangsúlyt az emelkedők, lejtők, valamint az ezek közötti átmenetek pontos kezelésére helyeztük. Ennek oka kettős: egyrészt ezek a jellegzetes útszakaszok erősen alulreprezentáltak az eddigi szakirodalomban és a nyilvánosan elérhető benchmarkokban, másrészt ha egy folytonosan és dinamikusan változó környezetben

kívánjuk előre jelezni az útfelület normálvektorát, elengedhetetlen az említett tényezők figyelembevétele. Az általunk kifejlesztett algoritmus tervezése és optimalizálása során ezért kiemelt szempont volt az emelkedési és lejtési szögek, valamint a hirtelen váltások okozta orientációváltások robusztus kezelése. Módszerünk nem csupán a statikus, sík útfelületeken mutat stabil teljesítményt, hanem olyan változatos domborzati viszonyok mellett is, amelyek a valós közlekedési környezetben rendszeresen előfordulnak, és amelyekhez a jelenleg alkalmazott modellek többsége csak korlátozott mértékben tud alkalmazkodni.

Az algoritmus teljesítményének reprezentatív bemutatásához az értékelést olyan, a szakirodalomban is elterjedten használt, nyilvánosan hozzáférhető adatkészleteken végeztük, amelyek jól lefedik a különböző domborzati viszonyokat és útfelület-orientációs kihívásokat. Az egyik ilyen a PandaSet, amely változatos és gyakran hirtelen bekövetkező meredekségváltozásokat tartalmaz, így különösen alkalmas az emelkedők és lejtők közötti váltások kezelésének vizsgálatára. A másik a KITTI adatkészlet, az egyik leggyakrabban használt referencia a korábbi kutatásokban – annak ellenére, hogy a felvételeken viszonylag kevés a jelentős síkváltozás. Emellett saját adatgyűjtést is végeztünk, amelynek során kifejezetten az előbbieken ismertetett dinamikus meredekségváltozásokat fedtük le, ezáltal hozzájárulva egy új kiértékelési standard kialakításához az útfelület-normálvektor predikciója területén. Az így létrehozott adatkészletet GradeSet névre kereszteltük. A három adatkészlet együttes alkalmazása lehetővé tette, hogy módszerünk teljesítményét széleskörűen, különböző környezeti és domborzati viszonyok között értékeljük, valamint összevessük a szakirodalomban publikált megoldásokkal.

Az 1. táblázat eredményei jól szemléltetik az általunk javasolt algoritmus kiváló általánosító képességét. Teljesítménye közel azonos szinten marad mindhárom vizsgált adatkészleten, annak ellenére, hogy azok jellegükben és környezeti kihívásaikban jelentősen eltérnek egymástól. A PandaSet sűrű útfelfestésekkel, de viszonylag rövid szekvenciákkal rendelkezik. A KITTI adatkészlet hosszú, többnyire sík terepszakaszokat tartalmaz, míg

1. táblázat | Az algoritmus eredményei az említett teljes adatkészleteken (PandaSet, KITTI, GradeSet), összehasonlítva az előző state-of-the-art algoritmussal

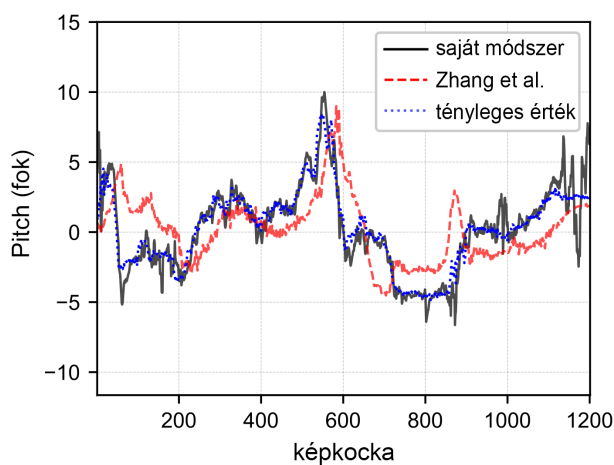
Adatkészlet/Módszer	RMSE (°)	AOE ₃ (%)	Lag (fr)	Normal error (deg)	Pitch error (deg)
PandaSet					
Zhang et al. 2022	0,83	2,91	4,45	1,68	0,62
Saját	0,55	1,24	0,30	0,61	0,38
KITTI (09–10 szekvenciák)					
Zhang et al. 2022	4,09	36,73	21,5	5,56	3,35
Saját	0,88	1,46	3,0	0,97	0,61
GradeSet					
Zhang et al. 2022	2,26	21,84	9,43	3,01	1,83
Saját	1,09	1,36	2,29	1,18	0,74

Forrás: Marko et al. 2026

a GradeSet kifejezetten széles meredekségtartományt ölel fel, dinamikus emelkedő/lejtő-váltásokkal. Kiemelendő, hogy minden mérési eredményt adatbázis-specifikus újrakalibrálás nélkül értünk el. A PandaSet, a KITTI és a GradeSet esetében is teljes mértékben azonos paraméter-beállításokat alkalmaztunk. Az ilyen változatos vezetési környezetekben tapasztalt, következetesen jó eredmények alátámasztják módszerünk robusztusságát és reprodukálhatóságát.

A következő ábrákon (2–4. ábra) minden adatkészlethez bemutatunk egy-egy pitch-hibagörbét, a leginkább kihívást jelentő emelkedő/lejtő-szakaszokat tartalmazó szekvenciák felhasználásával. A KITTI esetében a 10-es szekvenciát választottuk ki, amely a 2. ábrán látható pitch-görbe alapján mintegy 1200 képkockán keresztül körülbelül $+10^\circ$ és -5° között változik. Bár az átlagos lejtési meredekség kisebb, mint a másik két bemutatott példánál, a szekvencia jelentős nehézségeket tartogat: gyakori éles kanyarok, szűk útszakaszok miatti takarások, szembejövő forgalom, valamint parkoló járművek okozta látótérszűkülés. Az említett kedvezőtlen körülmények ellenére algoritmusunk ezen a szekvencián is stabil és pontos teljesítményt nyújtott, ami tovább erősíti a módszer gyakorlati alkalmazhatóságát valós, változó terepviszonyok között is.

Saját adatkészletünkéből a 25 százalékos emelkedőt tartalmazó szekvenciát választottuk ki (3. ábra), amely meredekségével jól szemlélteti az algoritmus teljesítményét és robusztusságát hirtelen emelkedések esetén. Megfigyelhető továbbá, hogy az összehasonlításként szereplő referenciaeljárás ezen a meredek szakaszon jelentős driftet mutat, miközben a mi megközelítésünk viszonylag stabil marad a teljes vizsgált időszak alatt. Ez tovább erősíti azt az állításunkat, miszerint a drift kritikus probléma a nagy emelkedőkkel tarkított útszakaszokon, és kiemeli módszerünk stabilitását ilyen körülmények között is. Fontos megjegyezni, hogy ebben a szekvenciában gyakorlatilag nem található felfestések

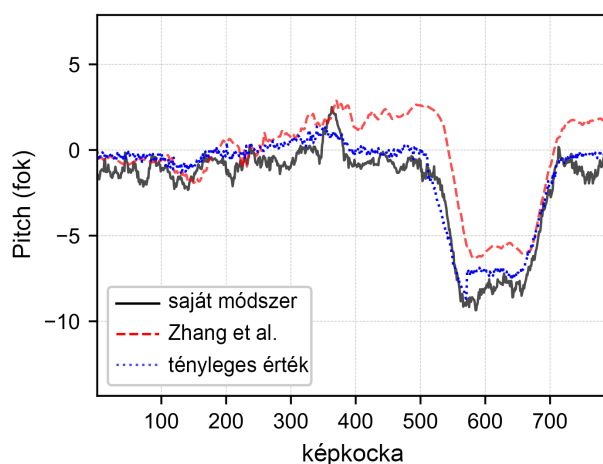


2. ábra | Saját eredményeink összehasonlítása a pitch-értékek vonatkozásában a KITTI adatkészlet 10. szekvenciáján Zhang et al. 2022 eredményeivel
Forrás: Markó et al. 2026

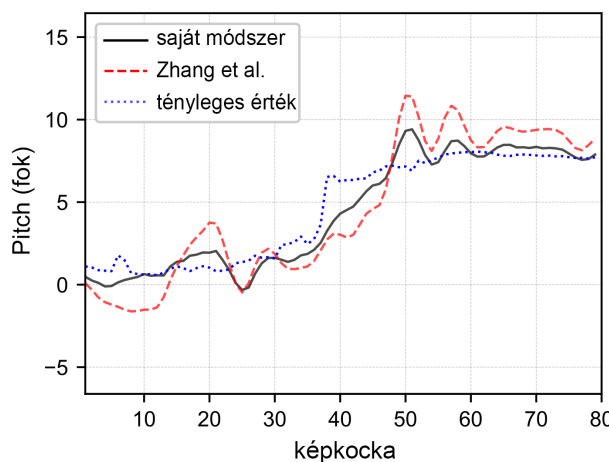
a jármű előtti szakaszon, ami jól demonstrálja a transzformeralapú matcher hosszú távú stabilitását.

A PandaSet az előző két adatkészlethez képest lényegesen rövidebb szekvenciákat tartalmaz. A 4. ábrán bemutatott példa egy összetettebb útemelkedést rögzít városi környezetben. A kedvezőtlen körülmények ellenére módszerünk stabil teljesítményt mutatott, a lejtőn való átmenetek során sima változással, valamint a referenciához képest kisebb túllövésekkel. Algoritmusunk átlagos normálhibája $1,18^\circ$, a pitch-komponens hibája pedig $0,75^\circ$ – szemben a referenciaeljárás $3,15^\circ$ -os és $1,72^\circ$ -os hibáival.

Összességében elmondható, hogy a kihívást jelentő jellemzők, az alacsony textúrájú útfelület, a meredek emelkedők és lejtők, valamint a hirtelen dőlésszög-változások ellenére az algoritmus képes robusztus útfelület-normálbecslést végezni különböző dinamikus scénáriókban is.



3. ábra | Saját eredményeink összehasonlítása a pitch-értékek vonatkozásában a GradeSet (saját) adatkészlet 11. szekvenciáján Zhang et al. 2022 eredményeivel
Forrás: Markó et al. 2026



4. ábra | Saját eredményeink összehasonlítása a pitch-értékek vonatkozásában a PandaSet adatkészlet 29. szekvenciáján Zhang et al. 2022 eredményeivel
Forrás: Markó et al. 2026

Diszkusszió

Ebben a fejezetben néhány kulcsfontosságú meglátást szeretnénk megosztani az eredmények alapján. Először is, a kísérleti eredmények azt mutatják, hogy az általunk javasolt, mesterségesintelligencia- és transzformerarchitektúra-alapú matcher jelentős mértékben javítja a homográfián alapuló talajsíkbecslés megbízhatóságát. Habár az egymást követő képkockák közötti homográfia számítása elméletben kézenfekvő megoldást kínál az útjelület síkjának becslésére, a gyakorlatban a jelleg-matching megbízhatatlansága miatt korlátozott körülmények között alkalmazható megbízhatóan.

Ez különösen igaz olyan esetekben, amikor a vizuális környezet textúraszegény, vagy ismétlődő mintázatokat tartalmaz (például úttest). Ezekben az esetekben a naív matching gyakran meghiúsul, ami nagymértékű becslési ingadozást eredményez, és ez túlkompenzálásra kényszeríti a szűrőket. Éppen ezt a stabilitásproblémát kezeli hatékonyan az általunk javasolt megközelítés. Az általunk alkalmazott matcher még a gyéren textúrázott útjelületeken is kiváló minőségű megfeleltetéseket biztosít. Ez jelentősen megkönnyíti a szűrés feladatát is, amely a rendszer stabilitásának másik forrása.

A SLERP-alapú szűrés lényegesen egyszerűbb, mint ha a Kálmán-szűrővel dolgoztunk volna, mivel nem igényel kovarianciamátrix-frissítéseket vagy mátrixinverziókat, ugyanakkor biztosítja az egységösszes kvaterniók fenntartását és az állandó szögsebességű forgást. Ezek a belső matematikai korlátok jól illeszkednek az útjelületek jellemzően finoman változó dőléséhez, még meredekebb lejtők esetén is. Ennek eredményeként a predikciók időben sokkal stabilabbak, miközben a számítási igény alacsony marad. Összességében a matcher általánosítható alkalmazhatósága új lehetőségeket nyit meg a normálvektor-becslési algoritmus kiterjesztése előtt más területeken is – ezeket a következőkben szeretnénk vázolni.

A kifejlesztett algoritmus egyik legkézenfekvőbb és azonnal hasznosítható felhasználási területe az adaptív sebességszabályozás. A jármű és az útjelület egymáshoz viszonyított orientációjának pontos ismerete lehetővé teszi, hogy a rendszer valós időben felmérje a lejtők, az emelkedők és az úthibák mértékét, és ezekhez igazítsa a jármű haladási sebességét. Ez különösen fontos lehet nagy tömegű, nehezen manőverezhető szárazföldi járművek, például nehéz szállítójárművek vagy katonai páncélozott járművek esetében, ahol a túlzott sebesség vagy a hirtelen lassítás a jármű stabilitásának romlásához, hosszabb fékúthoz vagy akár balesetveszélyhez vezethet. Védelmi alkalmazásokban a terepviszonyokhoz igazított sebesség nemcsak a biztonságot növeli, hanem a járművek rejtettségét és manőverezhetőségét is javíthatja, például akadályok kikerülése vagy a terepen való gyors helyváltoztatás során. Utasszállítási feladatokban pedig hozzájárulhat a járművek élettartamának növeléséhez, az energiafelhasználás optimalizálásához és az utazás komfortjának javításához.

Egy kevésbé nyilvánvaló, a kameraalapú autonóm rendszerekben ugyanakkor kulcsfontosságú gyakorlati felhasználási lehetőség – amelyet kutatásunkban is aktívan vizsgálunk – az útjelület prediktált normálvektorának alkalmazása a kameraszensor extrinzikus (külső) transzformációs mátrixának folyamatos kompenzálására az útjelület koordináta-rendszeréhez viszonyítva. Ennek szükségessége abból fakad, hogy a jármű és az út egymáshoz viszonyított helyzete a valós közlekedési környezetben folyamatosan változik: a kanyarodás, a fékezés, a gyorsítás, a felfüggesztés mozgása és az út egyenletlenségei mind befolyásolják a jármű dőlését vagy bólintását. Ezek a változások közvetlenül kihatnak a kamera és az útjelület közötti pillanatnyi orientációra, ami torzítja a képi mérésekből számított metrikus értékeket, például a pixelszintű távolságbecsléseket és a síkbeli projekciókat.

A szakirodalomban gyakran alkalmazott megoldás az úgynevezett „mount” (kamera–jármű) orientáció statikus meghatározása, például az eltűnési pont (vanishing point, VP) módszerrel, amely egyenes, sík úton, jól látható sáv-felfestések mellett képes megbízhatóan meghatározni a kamera yaw- és pitch-komponenseit a jármű koordináta-rendszerében. Ugyanakkor a VP önmagában nem alkalmas a gyors, valós idejű változások követésére, és érzékeny a sávjelzések hiányára vagy torzulására. Kutatásunkban ezt a hiányosságot hidaljuk át azzal, hogy mélytanuláson alapuló képfeldolgozási modellünk minden képkockában megbecsüli az útjelület normálvektorát a kamera koordináta-rendszerében, amelyet a kezdeti VP-alapú kalibrációval kombinálva valós időben frissítjük a teljes kamera–út transzformációt. Az útjelület normáljának folyamatos predikciója révén a kombinált transzformációs mátrix pontosan nyomon követhető, így a teljes transzformáció frissítésével jelentősen csökkenthető a hosszú távú drift, növelhető a pixelszintű mérésekből származó metrikus koordináták pontossága, és biztosítható a rendszer robusztus működése különböző út- és menetdinamikai körülmények között is. Ez a megközelítés lehetővé teszi, hogy a kamera–út kapcsolat a gyakorlatban is stabil és megbízható maradjon, ezáltal pontosabb és konzisztens érzékelést biztosítva autonóm járművek számára.

Amennyiben a fent bemutatott transzformációs mátrixot pontosan be tudjuk állítani, és folyamatosan tudjuk frissíteni, lehetőség nyílik arra, hogy pusztán kameraképek alapján egy háromdimenziós percepció környezetet hozzunk létre, amelyben leképezhetjük a detektált járható útjelületet, illetve a detektált objektumokat. Ez a megközelítés jelentős előnyt biztosíthat a számítási kapacitás szempontjából, mivel a vizuális alapú feldolgozás megfelelően optimalizált algoritmusokkal jóval kevesebb erőforrást igényel, ezáltal energiatöbbet, mint a nagyobb sávzélességet igénylő szenzorok, például a LiDAR feldolgozása. További előnye, hogy az alkalmazott kamerák általában jóval olcsóbbak, kisebbek, és elhelyezésük, rögzítésük is sokkal rugalmasabb, ami különösen előnyös lehet kisebb méretű, korlátozott teherbírású UGV-k (Unmanned Ground Vehicle) esetében.

A LiDAR-ral szemben a kamerák akár védettebb helyre is telepíthetők, kevésbé sérülékenyek, és tömegük, illetve energiaigényük is alacsonyabb, miközben a megfelelő feldolgozási láncsal képesek pontos, metrikusan értelmezhető 3D környezeti reprezentációt nyújtani a jármű autonóm navigációjához.

Algoritmusunk további fejlesztésekkel adaptálható lehet nem strukturált, nem sík terepen való alkalmazásra is, ahol a vizuális jellemzők térbeli és texturális eloszlása jelentősen eltér az épített utaktól. Ilyen off-road környezetekben a felszín heterogén mintázatai, a változó dőlésszögek, valamint a természetes akadályok (pl. sziklák, növényzet, vízfolyások) sokkal összetettebbé teszik az útfelület orientációjának és a járható területeknek a pontos meghatározását. Kutatásunk kezdetén valójában ebben az irányban indultunk el, azonban hamar nyilvánvalóvá vált, hogy a terepi viszonyok komplexitása, a vizuális változatosság kezelése, valamint a megbízható tanítóadatok előállítása első lépésként túl komplex lett volna. Ezért a kezdeti fejlesztési fókusz a strukturált, épített úthálózatokra helyeztük, ahol a sávfelfestések, a konzisztens felületi textúra és a kisebb topográfiai variancia lehetővé tették az algoritmus stabil alapjainak kialakítását. A jövőbeli fejlesztések során módszerünket olyan robusztus vizuális jellemzőkinyerési és felszínmodellezési technikákkal egészíthetjük ki, amelyek révén nem strukturált, off-road környezetekben is alkalmazhatóvá válik.

Köszönetnyilvánítás

A tanulmány elkészítését a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal (NKFIH) támogatta a Kooperatív Doktori Program Nemzetvédelmi Alprogram 2021 (NVKDP-2021) ösztöndíjprogramon keresztül.



Irodalomjegyzék

Alvarez, J. M., Gevers, T., LeCun, Y. & Lopez, A. M. (2012) Road scene segmentation from a single image. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y. & Schmid, C. (eds.) European Conference on Computer Vision – ECCV 2012. Lecture Notes in Computer Science, Vol. 7578. Springer, Berlin–Heidelberg. pp. 376–389. https://doi.org/10.1007/978-3-642-33786-4_28

Barath, D., Matas, J. & Noskova, J. (2019) MAGSAC: Marginalizing sample consensus. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10189–10197. <https://doi.org/10.1109/CVPR.2019.01044>

Chen, X., Kundu, K., Zhu, Y., Berneshawi, A. G., Ma, H., Fidler, S. & Urtasun, R. (2015) 3D object proposals for accurate object class detection. Proceedings of the 29th Conference on Neural Information Processing Systems (NeurIPS). pp. 424–432.

Hoiem, D., Efron, A. A. & Hebert, M. (2007) Recovering surface layout from an image. International Journal of Computer Vision, Vol. 75. No. 1. pp. 151–172. <https://doi.org/10.1007/s11263-006-0031-y>

Irani, M. & Anandan, P. (1996) Parallax geometry of pairs of points for 3D scene analysis. In: Buxton, B. & Cipolla, R. (eds.) Computer Vision – ECCV '96. 4th European Conference on Computer Vision. Lecture Notes in Computer Science, Vol. 1064. Springer, Berlin–Heidelberg. pp. 17–30. <https://doi.org/10.1007/BFb0015520>

Jeong, J. & Kim, A. (2016) Adaptive inverse perspective mapping for lane map generation with SLAM. 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI) 2016. pp. 38–41.

Košecák, J. & Zhang, W. (2005) Extraction, matching, and pose recovery based on dominant rectangular structures. Computer Vision and Image Understanding, Vol. 100. No. 2–3. pp. 274–293. <https://doi.org/10.1016/j.cviu.2005.04.005>

Markó, N., Rozsa, Z., Ballagi, A. & Sziranyi, T. (2026) Monocular ground normal prediction for the road ahead. IEEE Open Journal of Vehicular Technology, Vol. 7, pp. 1066–1080, <https://doi.org/10.1109/OJVT.2026.3676610>

Markó, N., Sziranyi, T. & Ballagi, A. (2023) Terrain depth estimation for improved inertial data prediction in autonomous navigation systems. Proceedings of the 2023 IEEE International Automated Vehicle Validation Conference (IAVVC). IEEE, Piscataway, NJ. pp. 1–6.

Pfeiffer, D. & Franke, U. (2010) Efficient representation of traffic scenes by means of dynamic stixels. IEEE Intelligent Vehicles Symposium (IV), La Jolla, CA, June 2010. IEEE. pp. 217–224.

Philion, J. & Fidler, S. (2020) Lift, splat, shoot: Encoding images from arbitrary camera rigs by implicitly unprojecting to 3D. Proceedings of the European Conference on Computer Vision. <https://doi.org/10.48550/arXiv.2008.05711>

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. & Zuiderveld, K. (1987) Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing, Vol. 39. No. 3. pp. 355–368. [https://doi.org/10.1016/S0734-189X\(87\)80186-X](https://doi.org/10.1016/S0734-189X(87)80186-X)

Reiher, L., Lampe, B. & Eckstein, L. (2020) A sim2real deep learning approach for the transformation of images from multiple vehicle-mounted cameras to a semantically segmented image in bird's eye view. IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC) 2020. pp. 1–7. <https://doi.org/10.48550/arXiv.2005.04078>

Shoemake, K. (1985) Animating rotation with quaternion curves. Proceedings of SIGGRAPH '85. Computer Graphics, Vol. 19. No. 3. pp. 245–254. <https://doi.org/10.1145/325165.325242>

Sun, J., Shen, Z., Wang, Y., Bao, H. & Zhou, X. (2021) LoFTR: Detector-free local feature matching with transformers. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2021, Nashville, TN. pp. 8922–8931. <https://doi.org/10.1109/CVPR46437.2021.00881>

Wang, Y., He, X., Peng, S., Tan, D. & Zhou, X. (2024) Efficient LoFTR: Semi-dense local feature matching with sparse-like speed. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2024. pp. 21666–21675. <https://doi.org/10.1109/CVPR52733.2024.02047>

Zhang, J., Sui, W., Zhang, Q., Chen, T. & Yang, C. (2022) Towards accurate ground plane normal estimation from ego-motion. Sensors, Vol. 22. No. 23. 9375. <https://www.mdpi.com/1424-8220/22/23/9375>

A cikk a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>) feltételei szerint publikált Open Access közlemény, melynek szellemében a cikk bármilyen médiumban szabadon felhasználható, megosztható és újraközölhető, feltéve, hogy az eredeti szerző és a közlés helye, illetve a CC License linkje és az esetlegesen végrehajtott módosítások feltüntetésre kerülnek. (SID_1)