

Dual-pixel CMOS szenzorstruktúra intra-frame mozgásdetekcióhoz



Németh Máté
A PhD Disszertáció tézisei

Pázmány Péter Katolikus Egyetem
Információs Technológiai és Bionikai Kar

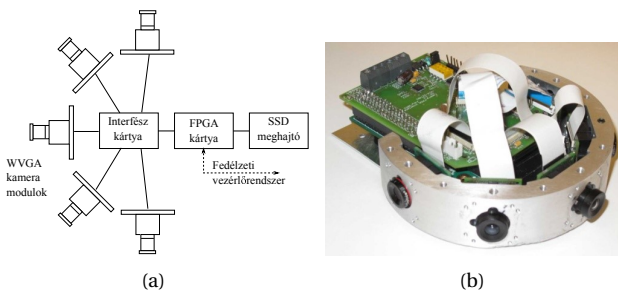
Témavezető:
Dr. Zarándy Ákos

Budapest, 2018

1. Bevezetés

Napjainkban a szenzorika egyik legdinamikusabban fejlődő ágát a képalakító szenzorok képezik. Ennek a fejlődésnek egy fontos területe a gépi látás, azon belül is a különböző autonóm járművek önálló feladatvégrehajtást támogató látórendszereivel kapcsolatos kutatások. Az MTA Számítástechnikai és Automatizálási Kutatóintézetében (SZTAKI) lehetőségem volt részt venni egy olyan projektben, melynek célja kisméretű pilótanélküli robotrepülőgépek (UAV) biztonságos önálló manőverezését biztosító fedélzeti rendszer kifejlesztése. A biztonságos autonóm műveletvégzés egyik fontos alapkövetelménye az ütközésselkerülés megvalósítása, melyhez egy többkamerás látórendszert építettünk [1].

Az 1. ábrán látható WVGA szenzorok képességeinek, hiányosságainak alaposabb vizsgálata indított el abban az irányban, hogy átfogó kutatást végezzek a napjainkban alkalmazott CMOS kamerák pixelstruktúrájával kapcsolatban. A képalakító szenzoroknak alapvetően két típusát különböztetjük meg: töltéscsatolt eszközök (CCD), illetve CMOS alapú kamerák. A CCD szenzorok érzékenysége a nagyobb kitöltési tényező miatt magasabb, és kevésbé zajérzékenyek, míg a CMOS eszközök legnagyobb előnye, hogy hasonló technológiával készülnek, mint a hagyományos VLSI áramkörök, így jóval egyszerűbb és költséghatékonyabb egy csipre integrálni különböző funkciókat. Ez lehetővé teszi a kutatók számára egyedi pixelstruktúrák megalkotását, és akár fókuszsíkbeli vagy szenzorközeli képfeldolgozási funkciók realizálását magán a szenzorcsipen.



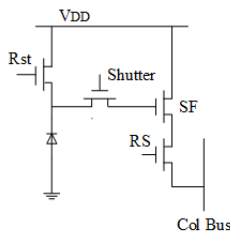
1. ábra. (a) A látórendszer blokkvázlata; (b) A komponensek összeszerelve

1.1. CMOS képalkotó szenzorok felépítése

A CMOS alapú szenzorok esetén a foton elektron konverzió több struktúrával történhet (fotodióda, fototranzisztor, fotogate). A legelterjedtebb fényérzékelési módszer a záró irányban előfeszített fotodióda feszültségesésének mérésén alapul. Az exponálás alatt a beeső fényintenzitástól függő fotoáram záróirányú áramként jelentkezik, kisütve a dióda pn átmenetének parazita kapacitását. A hullámhossztól függő fotoáram a következőképp számolható:

$$I_{ph} = \frac{q\eta P_{ph}}{\hbar\omega} \quad (1)$$

ahol I_{ph} a fotoáram, q az elemi töltés, η a kvantum hatásfok, P_{ph} a beeső fényteljesítmény, \hbar a redukált Planck állandó, és ω a körfrekvencia. Az exponálás végeztével, kiolvasáskor a fotodióda feszültsége hordozza a képi információt a pixelekben. Felépítését tekintve alapvetően két típusú pixelt különböztetünk meg: passzív, illetve aktív. A passzív pixelek a fotodiódán kívül csak egyetlen MOS kapcsolót tartalmaznak a kiolvasáshoz, ezzel szemben az aktív pixelekben egy aktív követőerősítő, és az exponálás vezérléséhez szükséges tranzisztorok is megtalálhatók. Az aktív pixel szenzorok (APS) a zár (shutter) vezérlési stratégiát tekintve két csoportra bonthatók. Az egyik csoportba az úgynevezett rolling shutter szenzorok tartoznak. Ebben az esetben sorról sorra végighaladva a szenzortömbön, az exponálás végeztével az adott sor kiolvasása azonnal megtörténik, így nincs szükség analóg tárolóelemekre a pixelekben. Emiatt ez a kiolvasási stratégia feltételezi, hogy az exponálás az egymást követő sorokban más és más időpillanatban kezdődik. Ezzel szemben a global shutter-t alkalmazó szenzorok (2. ábra) esetén koherens képet kapunk, mivel az exponálás minden pixel esetén azonos pillanatban kezdődik. Az áramköri megvalósítása viszont komplexebb az analóg tárolóelemek szükségessége miatt.



2. ábra. Egy klasszikus, négy tranzisztort tartalmazó global shutter aktív pixel struktúra (4T APS).

2. Módszerek

A látórendszer alapú sebességméréssel kapcsolatos tudományos eredmények alapvetően két kutatási irányhoz kapcsolódnak: optical-flow (inter-frame), és motion blur (intra-frame) alapú módszerek. Az inter-frame módszerek komoly irodalommal rendelkeznek, míg az intra-frame megoldásokkal kevés publikáció foglalkozik. Előbbi esetben az elmozdulás két képkocka (frame) összehasonlításából számítható. Ilyenkor a mérendő objektum detektálását mindkét képen el kell végezni, majd az objektumok megfeleltetését (matching) követően számolható azok elmozdulása a két kép között. Ez a feldolgozási szekvencia meglehetősen nagy számítási teljesítményt igényel.

Fentiekkel ellentétben az intra-frame mérési eljárások bemenete egyetlen kép, melyben az információt a motion blur effektus hordozza. A motion blur tulajdonképpen az objektum exponálás alatti elmozdulása miatt jön létre. A legtöbb esetben ez egy képminőséget rontó, nem kívánatos effektus, a mi esetünkben azonban a mérési eljárás alapját képezi. Hátránya, hogy amennyiben a mérendő objektum vizsgálatakor szükség van annak valamilyen azonosítására, klasszifikációjára, a kérdéses régiókból sokszor nehezen nyerhető ki használható képi információ. A 2. fejezetben bemutatott eljárás segítségével elérhető, hogy a motion blur csak a látvány bizonyos területein jelentkezzen szignifikánsan, míg a kép (vagy akár a mérendő objektum) többi része hordozza a további vizsgálatokhoz szükséges képi információt. Ezen felül megfelelő szenzorstruktúra alkalmazásával a mozgás-információt hordozó motion blur a szenzorban szétválasztható a látvány többi részétől, minimalizálva a szükséges utófeldolgozást, és ezáltal a számítási teljesítményt. Disszertációmban az intra-frame sebességmérés elméleti és gyakorlati vonatkozásait vizsgálom.

3. Új tudományos eredmények összefoglalása

Az új tudományos eredményeimet két téziscsoportba soroltam. Az első téziscsoportba az intra-frame sebességmérés elméletével kapcsolatos eredményeimet csoportosítottam, míg a második téziscsoportba az intra-frame mozgásdetekcióra optimalizált pixelstruktúrához kapcsolódó tézisek kerültek. A tudományos eredmények folyóiratokban ([J]-vel jelölve), illetve konferencia proceedings-ekben ([C]-vel jelölve) lettek publikálva.

1. Téziscsoport: Intra-frame mozgásdetekció és sebességmérés szuperponált képeken

A képkalkító szenzor felületére beeső fény mennyiség függ a látvány fényességétől (luminance, L_v), a lencse relatív apertúrájától (N), és az exponálási időtől (t). Egy mérési szituációban jellemzően N és L_v adott, így a mozgó objektumok intra-frame jellemzői, a motion blur hatása a képen t -vel szabályozható. Az intra-frame elmozdulásmérés során a mérési pontosságot a pixelizáció miatt elsősorban az exponálási idő határozza meg. Nagyobb t alkalmazásával nő a mérési pontosság, viszont a mozgó régiókban sok képi információ elveszik a motion blur miatt, tehát alapvetően egy trade-off jelentkezik a képminőség és a mérési pontosság között. Ebben a téziscsoportban módszert adtam fényforrással rendelkező objektumok intra-frame elmozdulásának mérésére.

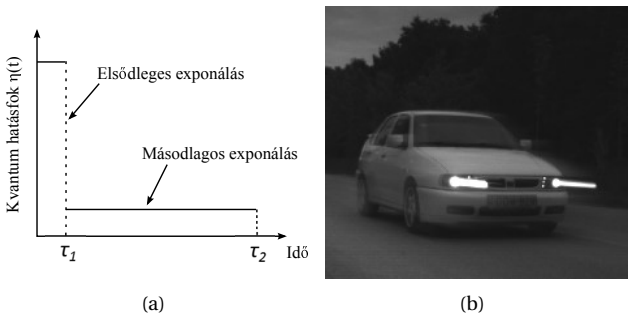
A téziscsoporthoz kapcsolódó publikációk: [J2], [C1]

1.1. Tézis Létrehoztam egy dupla exponáláson alapuló képrögzítési modellt képkalkító szenzorokhoz, ami lehetővé teszi szuperponált képek felvételét. Megmutattam, hogy alacsony shutter hatásfokú szenzorokkal emulálható ez az exponálási szekvencia.

A kifejlesztett módszer hatására létrejövő szuperponált kép két komponens összegeként áll elő. Az elsődleges kép tartalmazza a képi információt a látványról, míg a másodlagos képen csak a látvány nagy intenzitású területei jelennek meg, ez a komponens tartalmazza ezen területek intra-frame elmozdulását. Egy ilyen szuperponált kép létrehozásához kiegészítettem a klasszikus zár ciklust (nyitott, zárt) egy köztes - részben nyitott - állapottal. Az exponálási séma egyes fázisait különböző kvantum hatásfokkal modelleztem (1. ábra). Mivel a kvantum hatásfok a vezetési sávba gerjesztett töl-

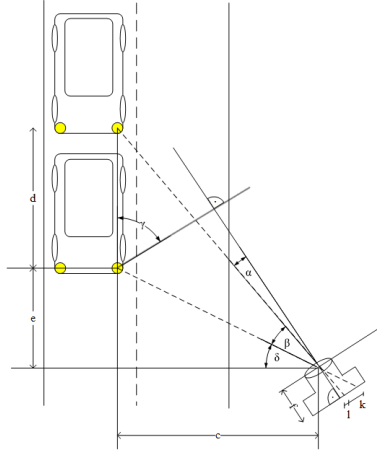
téshordozók számát adja meg az egyes beeső fotonokra, ezzel tulajdonképpen a szenzor érzékenységét változtattam az egyes fázisokban. A szuperponált kép a következőképpen jön létre: az elsődleges exponálás alatt rögzítésre kerül maga a látvány, majd a hosszabb másodlagos exponálás alatt, az alacsonyabb kvantum hatások miatt csak a nagy intenzitású régiókban módosulnak jelentős mértékben az előző fázis során eltárolt pixel értékek. Ezek alapján egy kellően nagy fényerejű mozgó fényforrás $[\tau_1, \tau_2]$ időintervallumhoz tartozó trajektóriája kirajzolódik a szuperponált képen, miközben az alacsony intenzitású területek, vagy a statikus fényforrások nem módosítják a rögzített képet.

Egy képalkotó szenzor shutter hatásfoka a pixel nyitott, illetve zárt shutter állapot melletti érzékenységének arányaként definiálható. Alacsony shutter hatásfoku szenzorok esetén tehát a zárt shutter állapotbeli töltésfelhalmozódás jelentősen módosíthatja a pixelben tárolt értéket, nagy intenzitású fényforrás akár szaturációba is viheti a pixelt annak kiolvasásáig. Ez a jelenség alkalmassá teszi az ilyen szenzorokat az intra-frame mozgásdetekcióra, ahogy az a 3. ábrán is látható.



3. ábra. (a) Az exponálási szekvencia sémája az elsődleges $[0, \tau_1]$ és a másodlagos $[\tau_1, \tau_2]$ exponálási fázisokkal; (b) Alacsony shutter hatásfokú szenzorral rögzített szuperponált felvétel

1.2. Tézis *A szenzorparaméterek és a mérési környezet geometriai paramétereinek ismeretében egy olyan új módszert dolgoztam ki, mely alkalmas szuperponált képeken a nagy intenzitású területek intra-frame elmozdulásának, illetve sebességének mérésére. Megvizsgáltam és mérési eredményekkel alátámasztottam az elérhető elméleti mérési pontosságot, ezen felül hibaanalízist végeztem, figyelembe véve az egyes paraméterek megváltozásának hatását a mérési pontosságra.*



4. ábra. A mérési környezet geometriája az alkalmazott jelölésekkel.

Amennyiben a mérési környezet geometriai paraméterei (4. ábra), és az exponálási szekvencia jellemzői ismertek, a látvány egy nagy intenzitású pontjának képsíkra vetített intra-frame elmozdulása alapján annak mozgási sebessége kifejezhető trigonometriai összefüggések segítségével.

$$d = c(\tan(\gamma - \alpha) - \tan(\gamma - \alpha - \beta)) \quad (2)$$

$$v = \frac{d}{\tau_2 - \tau_1} \quad (3)$$

Tekintve hogy a sebességmérés szögmérésre vezethető vissza, látható, hogy a sebességmérés hibája ekvivalens az intra-frame trajektória pixelben mért hosszának mérési hibájával. Egy konkrét mérési szituációban – ahol tehát adottak mind a geometria, mind a szenzor paraméterei – megvizsgáltam az elérhető mérési pontosságot egy mozgó jármű fényszórójának segítségével. Egy valós mérési elrendezés esetén – ahol a ground truth-t egy GPS-es sebességmérő szolgáltatta – 1,3%-os mérési hibát tapasztaltam.

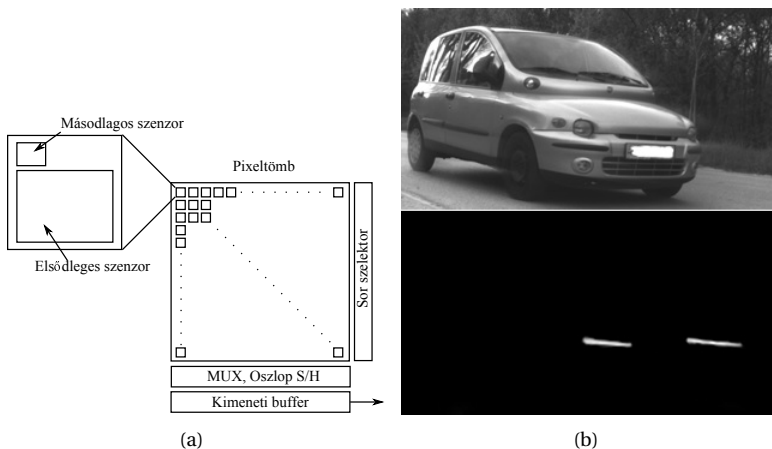
Amint az a 3. ábrán is látható, a szuperponált képeken végzett sebességmérés leginkább korlátozó tényezője a fényforrás trajektória kezdőpontja körüli szaturált régió, ami megnehezíti a fényforrás lokalizációját. Kidolgoztam egy két kamerán alapuló módszert, melynek segítségével sikerült megoldást találni a problémára.

2. Téziscsoport: Dual-pixel CMOS APS architektúra látórendszer alapú sebességméréshez

Az 1. Téziscsoportban bemutatott eljárás alapvető problémája, hogy a fényforrás lokalizálatlansága esetenként jelentős mérési hibát visz a rendszerbe. A 2. Téziscsoportban erre a problémára fókuszáltam, és kidolgoztam egy szilíciumon realizálható szenzorstruktúrát, ami a szenzorcsipen képes szeparálni a látvány nagy intenzitású régióinak intra-frame elmozdulását, emellett regisztrálni a látvány többi részét is.

A téziscsoporthoz kapcsolódó publikációk: [C2], [C3]

2.1. Tézis *Kidolgoztam egy pixel-szintű integrációs módszert látórendszer alapú sebességmérő szenzorokhoz, ami lehetővé teszi az 1.1 Tézisben definiált superponált kép egyes komponenseinek elkülönített rögzítését. Ezzel a kidolgozott új módszerrel tovább növeltem a sebességmérés pontosságát, a fényforrás trajektóriájának pontosabb lokalizálhatósága által.*



5. ábra. (a) A kidolgozott szenzor architektúra sematikus rajza a szubpixelekkel; (b) Egy példa a várt kimeneti képpárra.

Az intra-frame elmozdulásmérésre optimalizált dual-pixel szenzorstruktúra alapját az 1.1 Tézisben ismertetett dupla exponálási modell alkotja. A szenzor architektúráját úgy építettem fel, hogy az exponálási szekvencia egyes fázisaihoz külön szubpixelek tartozzanak. Minden egyes pixel tehát

kétféle szubpixelt tartalmaz: az elsődleges szubpixel valósítja meg az elsődleges exponálási fázist, míg a másodlagos szubpixelek a másodlagos fázist (5. ábra). Az ennek eredményeként létrejött struktúra két különálló pixeltömbje saját vezérlő bemenetekkel illetve adatkimenetekkel rendelkezik, viszont a rögzített képek geometriai és időbeli koherenciája miatt pixel szinten vannak integrálva. Ennek megfelelően, ahogy az 5. ábra is szemlélteti, a két független kimeneten két különböző exponálási paraméterekkel rögzített kép olvasható ki, így valósítva meg a nagy intenzitású régiók intra-frame elmozdulásának szenzor szintű szétválasztását a képi információt hordozó elsődleges képtől. A másodlagos pixeltömb érzékenységének beállításával biztosítható, hogy az általa rögzített képen csak a látvány magas intenzitású régiói jelenjenek meg.

2.2. Tézis *Kidolgoztam egy kvantitatív pixel maszkolási eljárást a 2.1 Tézisben bemutatott dual-pixel struktúrához, ami lehetővé teszi az intra-frame mozgásdetekciót végző subpixel érzékenységének beállítását a fotodióda aktív régiójának méretezésével.*

A módszer lényege, hogy amíg az elsődleges szenzor rögzíti a látványt, a másodlagos szenzor kimenetén, azonos fényáram mellett csak a nagy intenzitású területeken legyen nullától eltérő kimenet. Belátható, hogy ehhez nem szükséges abszolút skálán végrehajtani a méretezést, a két szubpixel érzékenységének aránya a meghatározó jellemző. A pixelek kimenetén megjelenő feszültség közelíthető az alábbi összefüggéssel:

$$V = \frac{A_{eff} t_{int}}{C_{jdep}} \int R(\lambda) f(\lambda) d\lambda \quad (4)$$

ahol A_{eff} a fotoaktív felület, t_{int} az exponálási idő, C_{jdep} a fotodióda pn átmenetének parazita kapacitása, $R(\lambda)$ a fotodióda érzékenysége a hullámhossz függvényében, míg $f(\lambda)$ a megvilágítás spektrális teljesítményeloszlás függvénye. Látható, hogy a technológia paramétereinek ismeretében a pixel-válasz a fotoaktív felület méretével befolyásolható. Ezt a jelenséget használtam ki a másodlagos szubpixel méretezése során.

4. Alkalmazási területek

A tézisekben megfogalmazott intra-frame sebességmérés egyik kézenfekvő alkalmazása járművek sebességének mérése. Napjainkban sebességmérésre leggyakrabban aktív érzékelőket alkalmaznak, mint például a radar és a lidar alapú rendszerek. Ezek drága, nagy fogyasztású eszközök, mindemellett a járművek azonosítása miatt a rendszerbe szükség van egy kamera integrálására is. Fentiek alapján jelentős előnnyel kecsegtetne egy olyan tisztán látórendszer alapú (tehát passzív) eszköz kifejlesztése, ami megfelelő pontossággal képes egyidejűleg a sebességmérésre és a járművek azonosítására.

A kifejlesztett két kamerás intra-frame sebességmérési eljárás, a hibaanalízis alapján alkalmasnak mutatkozik akár szigorú előírások teljesítésére is, azonban ennek alátámasztása további vizsgálatokat, validációt igényel. Az egy kamerás mérési módszer, a bemutatott korrekciós eljárás után kevésbé kritikus alkalmazásokban lehet megfelelő választás. Ilyen kevésbé szigorú specifikációjú alkalmazási terület lehet például egy elosztott forgalomellenőrző és irányító rendszer egy szenzor node-jaként történő alkalmazás. Mivel kisméretű, alacsony fogyasztású, olcsó eszközzel van szó, kézenfekvő megoldás lehetne a Smart City koncepcióban, akár nagy számban is alkalmazni a traffic management rendszerekben.

A szerző folyóirat publikációi

[J1] Ákos Zarándy, **Máté Németh**, Zoltán Nagy, András Kiss, Levente Sántha, Tamás Zsedrovits „A Real-Time Multi-Camera Vision System for UAV Collision Warning and Navigation,” *Journal of Real-Time Image Processing*, vol. 12, no. 4, pp. 709–724, 2016.

[J2] **Mate Nemeth**, Akos Zarandy „Intra-frame Scene Capturing and Speed Measurement Based on Superimposed Image: New Sensor Concept for Vehicle Speed Measurement,” *Journal of Sensors*, vol. 2016, Article ID 8696702, 10 pages, 2016. doi:10.1155/2016/8696702

[J3] Tamás Zsedrovits and Péter Bauer and Antal Hiba and Máté Németh and Borbála Jani Mátyásné Pencz and Ákos Zarándy and Bálint Vanek and József Bokor „Performance Analysis of Camera Rotation Estimation Algorithms in Multi-Sensor Fusion for Unmanned Aircraft Attitude Estimation,” *JOURNAL OF INTELLIGENT & ROBOTIC SYSTEMS*, vol. 84, no. 1-4, pp. 759–777, 2016

[J4] Tamás Zsedrovits and Péter Bauer and Borbála Jani Mátyásné Pencz and Antal Hiba and István Gőzse and Máté Németh and Zoltán Nagy and Bálint Vanek and Ákos Zarándy and József Bokor „Onboard visual sense and avoid system for small aircraft,” *IEEE AEROSPACE AND ELECTRONIC SYSTEMS MAGAZINE*, vol. 31, no. 9, pp. 18–27, 2016

A szerző konferencia publikációi

[C1] **M. Nemeth**, A. Zarandy „New sensor concept for intra-frame scene and speed capturing,” *European Conference on Circuit Theory and Design (ECCTD)*, Trondheim, 2015.

[C2] **M. Nemeth**, A. Zarandy, P. Földesy „Dual-pixel CMOS APS architecture for intra-frame speed measurement,” in *informal Proceedings of the IEEE International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS)*, Kosice, 2016.

[C3] **M. Németh**, Á. Zarándy, P. Földesy „Pixel-level APS Sensor Integration and Sensitivity Scaling for Vision Based Speed Measurement,” *30th anniversary Eurosensors conference, appears in Procedia Engineering*, vol. 168, pp. 1321–1324, 2016.

[C4] Tamás Zsedrovits and Péter Bauer and Máté Németh and Borbála Jani Mátyásné Pencz and Ákos Zarándy and Bálint Vanek and József Bokor „Per-

formance Analysis of Camera Rotation Estimation Algorithms for UAS Sense and Avoid,” *2015 Workshop on Research, Education and Development of Unmanned Aerial Systems, RED UAS 2015*, pp. 1–10, 2015.

[C5] Ákos Zarándy and Máté Németh and Borbála Jani Mátyásné Pencz and Zoltán Nagy and Tamás Zsedrovits „Cellular sensor-processor array based visual collision warning sensor,” *IEEE International Symposium on Circuits and Systems, ISCAS 2015*, pp. 1973–1976, 2015.

[C6] Ákos Zarándy and Tamás Zsedrovits and Borbála Jani Mátyásné Pencz and Máté Németh and Bálint Vanek „A Novel Algorithm for Distant Aircraft Detection,” *International Conference on Unmanned Aircraft Systems (ICUAS 15)*, pp. 774–783, 2015.

[C7] Tamás Zsedrovits and Ákos Zarándy and Borbála Jani Mátyásné Pencz and Antal Hiba and Máté Németh and Bálint Vanek „Distant aircraft detection in sense-and-avoid on kilo-processor architectures,” *2015 European Conference on Circuit Theory and Design*, pp. 1–4, 2015.

[C8] Ákos Zarándy and Zoltán Nagy and Bálint Vanek and T Zsedrovits and András Kiss and M Németh „A Five-Camera Vision System for UAV Visual Attitude Calculation and Collision Warning,” *Springer Lecture Notes in Computer Science*, no. 7963, pp. 11–20, 2013.

[C9] Ákos Zarándy and Zoltán Nagy and Tamás Zsedrovits and András Kiss and M Németh „FPGA implementation of a foveal image processing system for UAV applications,” *2014 14th International Workshop on Cellular Nanoscale Networks and their Applications - CNNA 2014*, pp. 1–2, 2014.

[C10] A Zarándy and B Pencz and M Németh and T Zsedrovits „Implementation of visual navigation algorithms on the Eye-RIS 1.3 system,” *2014 14th International Workshop on Cellular Nanoscale Networks and their Applications - CNNA 2014*, pp. 1474–1475, 2014.

[C11] Ákos Zarándy and Borbála Pencz and Máté Németh „Remote Aircraft Detection against Terrain Background and its implementation on SCAMP simulator,” *2014 14th International Workshop on Cellular Nanoscale Networks and their Applications - CNNA 2014*, pp. 1473–1474, 2014.