Report on the completion of OTKA NKFIH Research project 120499

Exploration of the environment in dynamic scenery from partial views

Tájékozódás dinamikus környezetben részleges látványokból

October, 2019

Tamas Sziranyi, SZTAKI

1. Year Plan

Goals:

- Initial set up of the project environment;
- Constructing the initial test databases;
- Building the single object library;
- Deep Learning: annotation, training, test;

Detection of objects and some vehicles;

- UAV imaging testing for registering with terrain sensors;
- Background and changing foreground separation for SLAM purposes;

Research methodology:

- Collecting test data to set up project environment by selecting test sites. Overview current research results, by using the participants' previous results as a starting point for research and publication goals. Preparing the Deep Learning database and testing the learning structure with the different data; Multi-time data fusion in DL CNN layers.

Tasks:

- Collect terrain, UAV and remote sensingdata for the test databases;
- Define test sites based on surveillance, UAV and remote sensing recommendation;
- Generate reference data for the test databases;
- Purchase requested software / hardware;
- Do the initial data processing: training and test for segmentation and recognition;
- Build the initial object library;
- Test existing features for describing single objects in training process;
- Testing initial data on Deep Learning structure
- Selection of features, fusion methods and learning structure
- Examining SLAM algorithms for implementing stochastic neighbourhood models.

Dissemination:

- Test laboratory and measurement sites setup
- Conference proceedings papers (2-3): IEEE and ISPRS Conferences and Workshops;
- Preparing and submitting 1-2 Journal publications.

Research results achieved in the first year:

In the work progress of this year we set up different measurement environments. We examined different SLAM algorithms for testing stochastic neighbourhood models. We have developed new algorithms for localizing, detection and visual interpretation. For the semantic analysis of environment we have developed and tested several deep learning and other methods.

In (Majdik et al. IJRR, 2017) it presents a dataset recorded on-board a camera-equipped Micro Aerial Vehicle (MAV) flying within the urban streets of Zurich, Switzerland, at low altitudes (i.e., 5-15 meters above the ground). The 2 km dataset consists of time synchronized aerial high-resolution images, GPS and IMU sensor data, ground-level street view images, and ground truth data. The dataset is ideal to evaluate and benchmark appearance-based localization, monocular visual odometry, simultaneous localization and mapping (SLAM), and online 3D reconstruction algorithms for MAVs in urban environments. This research has a high impact on our forthcoming work, where the SLAM methodology plays an important role.

In (Majdik et al, IWCIM 2016) we performed an extensive comparison and evaluation of the state-of-the-art online visual SLAM (Simultaneous Localization and Mapping) and offline visual SFM (Structure from Motion) methods in order to obtain the 3D model of the building. We show results obtained using a dataset recorded with a camera-equipped Micro Air Vehicle. This work ensures us to have the main SoA algorithms tested for our next purposes.

In (Manno-Kovacs, ICCV) we reported a monocular obstacle avoidance method based on a novel image feature map built by fusing robust saliency features, to be used in embedded systems on lightweight autonomous vehicles. The fused salient features are a textural-directional Harris based feature map and a relative focus feature map. We present the generation of the fused salient map, along with its application for obstacle avoidance. Evaluations are performed from a saliency point of view, and for the assessment of the method's applicability for obstacle avoidance in simulated environments. The presented results support the usability of the method in embedded systems on lightweight unmanned vehicles.

In different papers together with Domonkos Varga during 2016-2017 we have published results on the evaluation of semantic interpretation power of deep networks, namely Deep Convolutional Neural Networks (D-CNN). The evaluation metric was the color structure of recolored BW images, saying information about the deep learning power of complex structures in images. We have shown that pretrained (up to the level of semantic features) with additional training (from semantic to fine tuned coloring e.g.) methods can exploit sufficient information to prove the clustering efficiency.

In mapping the environment, a main task is to find the related objects around us. In our SMC-2016 paper we have shown a novel solution for a supervised learning framework that learns probability based semantic-level similarity and feature-level similarity simultaneously. This novel hashing scheme has a main advantage of that it is able to reduce the computational cost of retrieval significantly at the state-of-the-art efficiency level.

In our cooperated DFG project (not part of this project) our QoMEX paper is about crowd sourcing measure of video quality, what may have a great importance in our future research: namely the quality of the streamed video has a great influence on the SLAM and odometry measurement procedure later.

In our IJRS paper (not part of this project, but the basis to the forthcoming efforts) we have worked out the fusion of very different modalities, finding changes in a long term sense. Something similar is planned in our next approaches, when SLAM algorithms supply us the information about the possible positions of changes.

The environment detection needs sophisticated algorithms to find the main objects of high appearance probabilities, like pedestrians. In our preliminary work (D.Varga and T.Sziranyi, J AMBIENT INTELL HUMAN COMPUT) we detect pedestrians, considering applications including robotics, surveillance and automotive safety. We present a new feature extraction method based on Multi-scale Center-symmetric Local Binary Pattern operator. All the modules (foreground segmentation, feature pyramid, training, occlusion handling) of our proposed method are introduced with its details about design and implementation. Experiments on CAVIAR and other sequences show that the presented system can detect pedestrians in real-time effectively and accurately in surveillance videos.

Following this topic, in (D.Varga and T.Sziranyi, Eusipco 2017) a person re-identification based on Deep Multi-instance Learning has been shown for dynamic scene analysis.

Note: The two IWCIM conference papers (they are the preliminary work of this project) have been published during the period of NKFI grant, but have been submitted earlier, so NKFI grant ID could not be assigned.

2. Year Plan

Goals:

- Terrain cameras of mobile on-board and fixed surveillance cameras are registered together;
- Testing terrain SLAM models in dynamic situations: changes and estimation errors.
- UAV – ground devices cooperation: fusion of fixed and mobile views to form dynamic SLAM ;
- Improved feature extraction and single object detection;
- Scene description features;
- Air - ground camera view registration from video (co-motion in air - ground scenes)
- Detection of moving objects from moving cameras
- Change detection from comparing local maps and SLAM models;
- Semantic visual SLAM through training and recognition of objects and scenery events;

Research methodology:

- Based on the results of the existing feature extraction methods, features will be searched or novel ones will be developed. Features targeting reduced/eliminated mission parameter dependency (thus increasing "object-only" dependency) are selected to represent single objects and well defined scenery clusters in the object database. Test data will be classified to localize objects and scene situations. The database will be adapted for multiple data-sets to test site-independency.

Ground based (fixed terrain and mobile) cameras, also aerial UAV cameras are linked in common data processing.

Research results achieved in the second year:

In the second period we dealt with algorithms for SLAM optimization and pose-graph calculus, semantic image description based on deep learning, registration/calibration/fusion tasks for different modalities and/or very different views. We also investigated the timing/positioning/transmission problems for mobile sensor networks where we consider onboard vehicle- sensors, and made analysis on the mobile networks as image-based network registration background methods.

Some of the new achievements are basically engineering tasks to support our research, some mathematical methods like graph-based algorithms give the background to effective fusion and SLAM algorithms, while registration problems of multimodal fusion are a key question also in our research, but similar approaches are also used in other applications like

https://www.sztaki.hu/innovacio/hirek/taverzekeles-egrol-es-foldrol-fejlesztesunk-az-indexen.

In the second year, we have developed new algorithms to

- Recognition of partially viewed objects scanned by one-plane Lidar,
- Recognition of far-away objects by rare Lidar planes,
- Detecting depth-map from single image view by using deep learning,
- Detecting the malfunctions in a sensor-network by graph-analysis,
- Registering cameras on the terrain in the air for image fusion and multiview recognition,
- Recognizing objects on pixel-level semantic categorization by using deep learning architectures,
- New epipolar geometry calculus models have been constructed by using affine or rotation invariant correspondences,
- Optimizing the initial structure of pose graph used for Simultaneous Localization and Mapping (SLAM) systems.

In Harsanyi-Kiss-Majdik-Sziranyi_2018 [16] we have shown a new development of using a hybrid CNN approach for single image depth estimation.

One essential component of 3D video analysis is to develop faster and more reliable algorithms being capable of predicting depths from RGB images. Generally, it is easier to install a system with fewer cameras because it requires less calibration. Thus, our aim is to develop a strategy for predicting the depth on a single image as precisely as possible from one point of view. There are existing methods for this problem with promising results. The goal of this work is to advance the state-of-the-art in the field of single image depth prediction using convolutional neural networks. In order to do so, we modified an existing deep neural network to get improved results. The proposed architecture contains additional side-to-side connections between the encoding and decoding branches.

In Harsanyi-Kiss-Sziranyi_2018 [17] we give a solution to detect wormhole attack/malfunction against a sensor-network.

Wireless sensor networks and ad-hoc networks are gaining popularity rapidly due to their ability to solve challenging problems and the fact that thanks to recent technological advancements it is now possible to build smarter and denser networks. For example, they serve as the basis of the Internet of Things. Naturally, it is in the users best interest to develop increasingly secure networks. In some cases, the sensors are used in unknown or hostile environments. This and the vulnerability of the wireless communication channels used by arbitrary mobile communication networks means that they are exposed to various kinds of attacks or malfunctions. One of the most severe threats is the wormhole attack because an adversary can start the attack without compromising sensors or breaking through cryptographic defense mechanisms. In this paper, we propose a novel method for detecting wormhole attacks and identifying the affected sensors. Our approach does not rely on using any special measurement, only the connectivity information of the network.

In Harsanyi-Kiss-Sziranyi-Majdik_KEPAF [5, 27] we propose a novel algorithm to compute the initial structure of pose graph based Simultaneous Localization and Mapping (SLAM) systems.

We perform a Breadth-First Search (BFS) on the graph in order to obtain multiple votes regarding the location of a certain robot position from all of its previously processed neighbors. Next, we define the initial location of a pose as the average of the multiple alternatives. By adopting the proposed initialization approach the number of iteration needed for optimization is significantly reduced while the computational complexity remains lightweight. The quantitative evaluation using five benchmark datasets show the advantages of the proposed method.
This work has also been submitted to ***Pattern Recognition Letters***, being just under review after the revised version.

In Rózsa-Szirányi-2018 [2, 19] we propose a solution for an obstacle categorization problem for partial point clouds without shape modeling.
The approach is tested for a known database, as well as for real-life scenarios. In case of AGVs, real-time run is provided by on-board computers of usual complexity. It can recognize objects from very partial clouds: It is the main contribution of our paper. As our best knowledge it is the first time to attempt the recognition of objects from such low information as we did. This allows us to integrate it with some segmentation method and mutually benefit from each other. Applying the method with a safety scanner - localization sensor system, early prediction can enhance security in large extent. Finally this paper has been published in the ***IEEE Tr. Intelligent Transportation Systems***.

Included in our research, ***Daniel Barath*** [14] proposes a method for estimating an approximate fundamental matrix from six rotation invariant feature correspondences exploiting their rotation components, e.g. provided by SIFT or ORB detectors.
In the works of Barath et al. [15] new epipolar geometry calculus models have been constructed by using two affine correspondences or six rotation invariant correspondences. The first one's applicability is demonstrated on two-view multi-motion fitting, i.e. finding multiple fundamental matrices simultaneously, and outlier rejection; The second one is practical, which need to be done real time. He tested the method on 203 publicly available real image pairs.

3. Year Plan

Goals:
- Semantic visual SLAM for objects and scenes in dynamic situations;
- Continuous street/road/environment mapping and localization refreshment for traffic/transportation/work-area support;
- The SLAM model is supplemented by scouting MAV on-board camera views and remote sensing data to organize a cooperative sensor networks for Field of Interest.
- Change detection estimated from dynamic SLAM models
- Research methodology:
- Connecting visual chains for generating closed SLAM loops from partial view sections;
- Remote Sensing data (slow time-scale) contributing to quick localization and mapping;
- SLAM and change detection algorithms for recognition from change maps.

Tasks:
- Evaluate different data-sets on the same area;
- Localize changes on SLAM results of different looping iterations;
- Looping sections of view series to build closed SLAM loops;

- Change detection estimated from dynamic SLAM models;
- Terrain, low-altitude and satellite data will be registered and matched in content,
- Object detection and scene interpretation methods, based on learning, will result in semantic evaluation.

Research results achieved in the third year:

In the last, third year we have completed our toolkit of methodologies by computer geometry solutions for better positioning and SLAM calculus for measuring repeatedly motion tracks and comparing them to find changes.

Localization of map changes by exploiting SLAM residuals [34]
Simultaneous Localization and Mapping is widespread in both robotics and autonomous driving. This paper proposes a novel method to identify changes in maps constructed by SLAM algorithms without feature-to-feature comparison. We use ICP-like algorithms to match frames and pose graph optimization to solve the SLAM problem. Finally, we analyze the residuals to localize possible alterations of the map. The concept was tested with 2D LIDAR SLAM problems in simulated and real-life cases.

Fusion Markov Random Field Image Segmentation for a Time Series of Remote Sensed Images [31]
Since outdoor sceneries, principally observation of natural reserves, agricultural meadows and forest areas, are changing in illumination, coloring, textures and shadows time-by-time, and the resolution and geometrical properties of the imaging conditions may be also diverse, robust and semantic level algorithms should be developed for the comparison of images of the same or similar places in very different times.
Earlier, a new method, fusion Markov Random Field (fMRF) method has been introduced which applied unsupervised or partly supervised clustering on a fused image series by using cross-layer similarity measure, followed by a multi-layer Markov Random Field segmentation. This work shows the effective parametrization of the fusion MRF segmentation method for the analysis of agricultural areas of fine details and difficult subclasses.

Object detection from a few LIDAR scanning planes, IEEE Tr. IV
LIDAR sensors enable object and free-space detection for intelligent transportation systems and vehicles. This work [28] proposes a recognition method for LIDARs based on only a few detection planes. This method is useful especially in the case when the angular resolution of the scan is sufficient, but in the vertical direction, the planes are far from each other. We use Fourier descriptor to characterize a scan plane and Convolutional Neural Network for classification. Our method exploits both time varying shape information and contours from multiple scan planes if available. The method performs at least as well as the state of the art algorithms in case of near field, and it also expands the detection range. It was evaluated on tens of thousands of samples from large public datasets and we did separate evaluation for far field objects as well. The stages of this work has also been published in [20, 28, 29, 30], including an IEEE conference and ***IEEE Tr. Intelligent Vehicles*** [28].

Optimal Multi-View Surface Normal Estimation Using Affine Correspondences
Barath et.al have introduced (IEEE TIP [25]) an optimal, in the least squares sense, method is proposed to estimate surface normals in both stereo and multi-view cases. The proposed algorithm exploits exclusively photometric information via affine correspondences and estimates the normal for each correspondence independently. The normal is obtained as a root

of a quartic polynomial. Therefore, the processing time is negligible. Eliminating the outliers, we propose a robust extension of the algorithm that combines maximum likelihood estimation and iteratively re-weighted least squares. The method has been validated on both synthetic and publicly available real-world datasets. It is superior to the state of the art in terms of accuracy and processing time. Besides, we demonstrate two possible applications: 1) using our algorithm as the seed-point generation step of patch-based multi-view stereo method, the obtained reconstruction is more accurate, and the error of the 3D points is reduced by 30% on average and 2) multi-plane fitting becomes more accurate applied to the resulting oriented point cloud. An important other result in similar topic has been published by Ivan Eichhardt in **BMVC** [26]: Optimal Multi-view Correction of Local Affine Frames. Another important result in Vomputer Geometry has been published by D. Barath on **ACCV** [24]: Recovering Affine Features from Orientation- and Scale-Invariant Ones.

It was not in the project's goals, but we have also worked on image- (with a DFG project in Kontsanz [6, 33]) and on video (with Domonkos Varga [32]) quality measurement method, since SLAM and similar positioning algorithms are dependent on it. We also investigated the image distortion and quality issues for the recognition in deep learning architecture. It will be important later, when image sampling and transmission will be a key issue in image-based localization tasks, like random source SLAM.

Summary of the work

In this work, we fused the image-like data from different modalities: camera, Lidar, infra, multi-channel. We used new solutions for data-fusion, recognition and calibration in 3D computer geometry. For the positioning and repeated SLAM tracking we have introduced a new fast and robust algorithm for pose-graph initialization, and a wormhole detection method in wireless sensor networks to find source-errors in the sensors'space. We also have developed a novel change detection solution from SLAM error-estimation procedure instead of 3D point cloud generation and comparison.

Deep learning has important role in several cases in the project to define semantic information: monocular depth map and object categorization, finding the most appropriate colors for color-less images as a semantic proof, recognition of objects of poor data-set, or blind image- or video-quality estimation. For the above works we had to generate some own image/video databases, and also using reference data-sets for correct comparison. The proof of the development in this competitive area is a serious demand, why a large part of our research activity had been devoted for careful comparison and evaluation discussions. However, we have achieved strong development in several cases, published them in the best journals and conferences.

In many cases, the sensorial information is incomplete or noisy; we had solutions for these cases to retrieve as much additional semantic meaning from the data as possible. In addition, the complex sensorial information helped us to use new fusion methods, geometry calculus and statistical evaluation.

General goals fulfilled

Overall outcome:

[1] Scientific results will be published in prestigious journals (IEEE, ISPRS) and on the most relevant conferences (ICCV, CVPR, ICPR, ICIP, ICRA, IROS). ***Most of them have been fulfilled.***

[2] The project will result in algorithmic toolkits for dynamic SLAM modules, dynamic change detection and stochastic fusion algorithms. ***The programs are free for tests.***

[3] Scientific and technical results, together with the experimental application and validation setup will be demonstrated and tested in the experimental demonstration system configured during the project. ***Most of the algorithms have been demonstrated in real-life situations***.

Comparison of the planned to the completed work

Problems have been solved:	**Related publications**
1. Object recognition and scene analysis by using Deep Leaning	12, 13, 16, 22, 23, 28, 29, 32, 33
2. Detection of changes and dynamic features with varying sensor network	
- Change detection in long-term basis	9, 31
- Change detection during refreshed SLAM calculus	34
3. Remote sensing (aerial and satellite) –co-registration by using dynamic descriptors	
4. Visual SLAM algorithms in dynamic environments	27
- New saliency features	6, 8
- descriptors for characterizing the foreground	11
5. SLAM algorithms based on terrestrial (fix and on-board mobile camera network) and remote sensing views	1, 7, 34
Tasks in general:	
1. Deep Learning: annotation, training, test Detection of objects: pedestrians, vehicles street furniture, buildings	3, 4, 6, 7, 10, 13, 16, 20, 30
2. Air-ground camera view registration from video	1, 7, 21
3. View registration btw remote sensing cameras and terrestrial cameras (new features for registration)	14, 15, 24, 25, 26
4. Air-ground event detection and tracking	8, 9, 28-30
5. Modality fusion: RGB, depth, LIDAR, multispectral (mainly using earlier development)	9
6. Detection of moving objects from moving cameras (mainly using earlier development)	2
7. Semantic visual SLAM	16
8. Air-ground collaborative SLAM in dynamic environment	5, 7
9. Connecting partial scanning sequences for building a closed loop for SLAM calculus	5, 34
10. Change detection estimated from dynamic SLAM models	34

Publications

[1] Majdik A L, Tizedes L, Bartus M, Szirányi T: Photogrammetric 3D reconstruction of the old slaughterhouse in budapest, In: IEEE (szerk.) (szerk.) 2016 International Workshop on Computational Intelligence for Multimedia Understanding (***IWCIM***). Piscataway: IEEE, 2016. pp. 1-5., 2016 *

[2] Rozsa Z, Sziranyi T: Object detection from partial view street data, In: IEEE (szerk.) (szerk.) 2016 International Workshop on Computational Intelligence for Multimedia Understanding (***IWCIM***). Piscataway: IEEE, 2016. pp. 1-5., 2016 *

[3] Varga D, Szirányi T: Fast content-based image retrieval using Convolutional Neural Network and hash function, In: Szakál A (szerk.) (szerk.) 2016 IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings: ***SMC*** 2016.: IEEE, 2016. pp. 2636-2640., 2016 *

[4] Varga D, Szirányi T: Fully automatic image colorization based on Convolutional Neural Network, In: IEEE (szerk.) (szerk.) 23rd International Conference on Pattern Recognition (***ICPR***). Cancun: IEEE, 2016. pp. 3680-3685., 2016 *

[5] Harsányi K, Kiss A, Majdik A, Szirányi T: Robusztus kezdeti pozíció-gráf előállítása hatékony szimultán helymeghatározási és térképezési rendszerekhez, NJSZT KÉPAF 2017: Képfeldolgozók és Alakfelism, 2017

[6] Hosu V, Hahn F, Jenadeleh M, Lin H, Men H, Szirányi T, Li S, Saupe D: The Konstanz natural video database (KoNViD-1k), In: [sn] (szerk.) (szerk.) 9th International Conference on Quality of Multimedia Experience, ***QoMEX*** 2017. [s. l.] - Nemzetközi: IEEE, 2017. pp. 1-6., 2017 *

[7] Majdik AL; Till C; Scaramuzza D: The Zurich urban micro aerial vehicle dataset, ***INTERNATIONAL JOURNAL OF ROBOTICS RESEARCH***, 2017

[8] Manno-Kovács Andrea, Kovács Levente: Lightweight Monocular Obstacle Avoidance by Salient Feature Fusion, ***ICCV*** 2017, IEEE, Velence, 2017

[9] Shadaydeh, M, Zlinszky, A, Manno-Kovács, A , Sziranyi,T: Wetland mapping by fusion of airborne laser scanning and multi-temporal multispectral satellite imagery, ***INT J REMOTE SENS*** 38: (23) 7422-7440, 2017

[10] Varga D, Szabó Cs A, Szirányi T: Automatic Cartoon Colorization Based on Convolutional Neural Network, In: ACM (szerk.) (szerk.) CBMI '17. Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing. New York: ACM Press, 2017. pp. 1-6., 2017

[11] Varga D, Szirányi T: Robust real-time pedestrian detection in surveillance videos, ***J AMBIENT INTELL HUMAN COMPUT*** 8: (1) 79-85, 2017

[12] Varga D, Szirányi T: Twin deep convolutional neural network for example-based image colorization, ***LECT NOTES ARTIF INT*** 10424: 184-195, 2017 *

[13] Varga D, Szirányi T: Person Re-identification based on Deep Multi-instance Learning, In: EURASIP (szerk.) (szerk.) 25th European Signal Processing Conference (***EUSIPCO***). pp. 1604-1608., 2017 *

[14] Barath D: Approximate Epipolar Geometry from Six Rotation Invariant Correspondences, Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5: VISAPP, 464-471, 2018, Funch, 2018

[15] Baráth D, Hajder L: Efficient Recovery of Essential Matrix From Two Affine Correspondences, ***IEEE Tr. Image Processing***, 2018

[16] Harsányi K, Kiss A, Majdik A, Sziranyi T: A Hybrid CNN Approach for Single Image Depth Estimation: A Case Study, Multimedia and Network Information Systems. Proceedings of the 11th International Conference MISSI 2018. Konferencia helye, ideje: Wrocław, Lengyelország, 2018

[17] Harsányi Károly, Kiss Attila, Szirányi Tamás: Wormhole Detection in Wireless Sensor Networks Using Spanning Trees, 2018 ***IEEE*** International Conference on Future IoT Technologies, Future IoT 2018., 2018

[18] Manno-Kovács A, Majdik A, Szirányi T: Satellite and Aerial Image Processing for Smart Farming and Biodiversity Conservation, ***ERCIM NEWS*** (113) pp. 33-34., 2018 *

[19] Rózsa Z, Szirányi T: Obstacle Prediction for Automated Guided Vehicles Based on Point Clouds Measured by a Tilted LIDAR Sensor, ***IEEE Transactions on Intelligent Transportation Systems***, 2018

[20] Rózsa Z, Szirányi T: Városi objektum felismerés mindösszesen néhány LIDAR szkennelési síkból, IX. magyar számítógépes grafika és geometria konferencia, GRAFGEO 2018., 2018

[21] Szirányi T, Havasi L: Observation on Earth and from the sky, In: Finszter, G; Sabjanics, I (szerk.) Security challenges in the 21st century, Dialóg Campus Kiadó (2018) pp. 401-410., 2018 *

[22] Varga D, Saupe D, Szirányi T: DeepRN: A Content Preserving Deep Architecture for Blind Image Quality Assessment, IEEE International Conference on Multimedia and Expo (***ICME***) 2018. San Diego, 2018

[23] Varga D, Szirányi T: A comparative study of about how image quality influences convolutional neural networks, IX. magyar számítógépes grafika és geometria konferencia, GRAFGEO 2018., 2018

[24] Daniel Barath: Recovering Affine Features from Orientation- and Scale-Invariant Ones, ACCV 2018: Computer Vision – ***ACCV*** 2018 pp 266-281, LNCS V. 11361, 2019

[25] Daniel Barath, Ivan Eichhardt, Levente Hajder: Optimal Multi-View Surface Normal Estimation Using Affine Correspondences, ***IEEE TRANSACTIONS ON IMAGE PROCESSING***, VOL. 28, NO. 7, JULY 2019, 2019

[26] Ivan Eichhardt,: Optimal Multi-view Correction of Local Affine Frames, 30th British Machine Vision Conference, ***BMVC*** 2019, pp. 1-12. Paper: 0816, 2019

[27] Karoly Harsanyi, Attila Kiss, Tamas Sziranyi, Andras Majdik: MASAT: a fast and robust algorithm for pose-graph initialization, ***Pattern Recognition Letters***, revision submitted, 2019

[28] Rozsa Z, Sziranyi T: Object detection from a few LIDAR scanning planes, ***IEEE Transactions on Intelligent Vehicles*** xx: (x) pp. 1-13., 2019 *

[29] Rozsa Z, Sziranyi T: Street object classification via LIDARs with only a single or a few layers, In: IEEE (szerk.) 3rd IEEE International Conference on Image Processing, Applications and Systems, IPAS 2018, ***IEEE*** (2019) pp. 156-161., 2019 *

[30] Rózsa Z, Szirányi T: Utcai objektumok osztályozása LIDAR adatokon csak egyetlen vagy néhány síkból, In: KÉPAF 2019. Képfeldolgozók és Alakfelismerők Társaságának 12. országos konferenciája, (2019) pp. 1-15., 2019 *

[31] Szirányi T, Kriston A, Majdik A, Tizedes L: Fusion Markov Random Field Image Segmentation for a Time Series of Remote Sensed Images, In: Faragó, I; Izsák, F; Simon, P L

(szerk.) Progress in Industrial Mathematics at ECMI 2018, 20th European Conference on Mathematics for Industry, Springer (2019) pp. 99-107., 2019 *

[32] Varga D, Szirányi T: No-reference video quality assessment via pretrained CNN and LSTM networks, ***SIGNAL IMAGE AND VIDEO PROCESSING*** pp. 1-8., 2019 *

[33] Vlad Hosu, Hanhe Lin, Tamas Sziranyi, Dietmar Saupe: KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment, ***IEEE TRANSACTIONS ON IMAGE PROCESSING***, AQ subm, 2019

[34] Zoltan Rozsa, Marcell Golarits , Tamas Sziranyi: Localization of map changes by exploiting SLAM residuals, ***ACIVS*** 2020 submitted, 2020