POSTER SESSION BOOKLET

International Computer Vision Summer School 2016

Computer Vision: What Happens Next?

Sicily, 17-23 July 2016

International Computer Vision Summer School

Computer Vision is the science and technology of making machines that see. It is concerned with the theory, design and implementation of algorithms that can automatically process visual data to recognize objects, track and recover their shape and spatial layout.

The International Computer Vision Summer School - ICVSS was established in 2007 to provide both an objective and clear overview and an in-depth analysis of the state-of-the-art research in Computer Vision. The courses are delivered by world renowned experts in the field, from both academia and industry, and cover both theoretical and practical aspects of real Computer Vision problems.

The school is organized every year by University of Cambridge (Computer Vision and Robotics Group) and University of Catania (Image Processing Lab). The general entry point for past and future ICVSS editions is:

<div align="center">

http://www.dmi.unict.it/icvss

</div>

ICVSS Poster Session

The International Computer Vision Summer School is especially aimed to provide a stimulating space for young researchers and Ph.D. Students. Participants have the possibility to present the results of their research, and to interact with their scientific peers, in a friendly and constructive environment.

This booklet contains the abstract of the posters accepted to ICVSS 2016.

Best Presentation Prize

A subset of the submitted posters will be selected by the school committee for short oral presentation. A best presentation prize (supported by Qualcomm, Facebook, Toshiba and NVIDIA) will be given to the best presentations selected by the school committee.

Scholarship

A scholarship (supported by Qualcomm, Facebook, Toshiba and NVIDIA) will be awarded to the best PhD student attending the school. The decision is made by the School Committee at the time of the School, taking into account candidates' CV, poster and oral presentation.

Sicily, June 2016   *Roberto Cipolla*
Sebastiano Battiato
Giovanni Maria Farinella

List of Posters[1]

1. LIGHTFIELD IMAGE PROCESSING AND PHOTOMETRIC STEREO
   Antensteiner D.

2. TISSUE-BACKGOUND SEGMENTATION ON HISTOPATHOLOGICAL WHOLE SLIDE IMAGES
   Bándi P., Ciompi F., Litjens G. J. S., Laak van der J. A. W. M.

3. TEXT LOCALIZATION IN THE WILD WITH FCN
   Bazazian D., Karatzas D., Bagdanov A.

4. RECONSTRUCTION VIA DETECTION: EFFICIENT AND AUTOMATIC RECONSTRUCTION FROM UNORGANIZED 3D SCANS
   Birdal T., Ilic S.

5. GLOBAL OPTIMALITY VERIFICATION IN POSE GRAPH OPTIMIZATION
   Briales J.

6. OBJECT RECONSTRUCTION EXPLOITING SEGMENTATION
   Bullinger S.

7. SHUKB: STUDY OF HUMAN KINETIC BEHAVIOR
   Burbano A., Bouaziz S., Vasiliu M.

8. TEMPORAL ACTIVITY PROPOSALS FOR EFFICIENT DETECTION OF HUMAN ACTIONS IN UNTRIMMED VIDEOS
   Caba Heilbron F., Niebles J.C., Ghanem B.D

9. PEOPLE COUNTING BY HUBER LOSS REGRESSION
   Cavazza J., Murino V.

10. MOTION ANALYSIS IN DYNAMIC TEXTURES
    Chubach O.

11. ENHANCING EARLY LEARNING IN IMAGE CLASSIFIERS WITH COMPLEX FILTERS
    Cotter F., Kingsbury N.

12. LEARNING TRANSFERABLE POLICIES FOR AUTONOMOUS MAV CONTROL
    Daftry S., Bagnell J. A., Hebert M.

13. DISCRIMINATIVE CORRELATION FILTERS FOR VISUAL TRACKING
    Danelljan M.

14. 2D IMAGE CLASSIFICATION FOR 3D ANATOMY LOCALIZATION
    De Vos B.D., Wolterink J.M., De Jong P.A., Viergever M.A., Išgum I.

---

[1]Posters are ordered by surname of first author. Each poster is identified by a number. The page of a poster in this booklet corresponds with the ID of the poster.

LIGHTFIELD IMAGE PROCESSING AND PHOTOMETRIC STEREO

Antensteiner D.

Abstract: We analyze light rays of objects in order to estimate their depth. Conventional stereo systems only capture two views of the object under constant illumination. We achieve a more complete description by using both lightfield imaging and photometric stereo techniques. Lightfields describe the variation of the image over observational directions. We will use their information to obtain a 3D reconstruction of the scene. Photometric stereo describes the variation of surfaces in respect to the lighting direction. We fuse the information of both disciplines to gain a refined depth map.

Contact: doris.antensteiner.fl@ait.ac.at

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

TISSUE-BACKGOUND SEGMENTATION ON HISTOPATHOLOGICAL WHOLE SLIDE IMAGES

Bándi P., Ciompi F., Litjens G. J. S., Laak van der J. A. W. M.

Abstract: In this poster we propose a method to automatically segment the tissue and background in histopathological whole slide images using deep learning.

Contact: peter.bandi@gmail.com

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

TEXT LOCALIZATION IN THE WILD WITH FCN

Bazazian D., Karatzas D., Bagdanov A.

Abstract: Reading text in the wild has gained plenty of attention in the computer vision community. We investigate how latest results on deep learning can advance text localization pipelines. Fully convolutional networks (FCNs) and derivative methods have achieved great performance on pixel level recognition tasks. We train an FCN to generate the probability of each pixel in the original image being part of a text region. Text localization is performed by thresholding the normalised prediction-map of the FCN.

Contact: dena.bazazian@gmail.com

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

RECONSTRUCTION VIA DETECTION: EFFICIENT AND AUTOMATIC RECONSTRUCTION FROM UNORGANIZED 3D SCANS

Birdal T., Ilic S.

Abstract: We propose reconstruction-via-detection framework, a radically new and effective approach to 3D reconstruction from unordered and unstructured scans, containing clutter and occlusions. Our pipeline uses a nominal CAD model to guide the reconstruction. First, the model is matched to the scenes, finding initial alignment and segmentation. Then we construct a pose graph and globally align the segmented scans. Our method is applicable to reconstructions of objects from small to very large volumes.

Contact: tolga.birdal@tum.de

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

GLOBAL OPTIMALITY VERIFICATION IN POSE GRAPH OPTIMIZATION

Briales J.

Abstract: Graph-based SLAM is one of the most effective solutions to the Simultaneous Localization and Mapping problem. This framework leads to the optimization of a high-dimensional, non-convex problem. Thus, state-of-the-art local optimization techniques may get stuck in local minima arbitrarily far from the globally optimal point. In our work we meet this shortfall by leveraging well-established results from duality theory to certify global optimality of potential solutions.

Contact: jesusbriales@gmail.com

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

OBJECT RECONSTRUCTION EXPLOITING SEGMENTATION

Bullinger S.

Abstract: We present a method to reconstruct the 3D shape of a possibly moving object in video data. Our approach is robust to occlusion, handles stationary objects and extends naturally to multiple video sequences. Our algorithm allows us to compute object meshes by leveraging semantic segmentations and virtual camera positions. We evaluate our method using a multi-view dataset of a moving car. A laser scan of the vehicle serves as ground truth. The algorithm achieves an average point distance of 3.3 cm.

Contact: sebastian.bullinger@iosb.fraunhofer.de

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

SHUKB: STUDY OF HUMAN KINETIC BEHAVIOR

A. Burbano, S. Bouaziz, M. Vasiliu

Abstract: Smart environments require real-time tracking on large public and private spaces. Usually, these environments are plenty of physical barriers generating occlusion. Due to the huge coverage and obstacles, several cameras are required. The real-time people characteristics and behavior detection increment the required computing level. We create a scalable distributed smart camera network dedicated to track and count people. In the network, each node is capable of both tracking people inside of a global coordinate system and communicating tracking information with adjacent nodes

Contact: andres@shopline.fr

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

TEMPORAL ACTIVITY PROPOSALS FOR EFFICIENT DETECTION OF HUMAN ACTIONS IN UNTRIMMED VIDEOS

Caba Heilbron F., Niebles J.C., Ghanem B.D

Abstract: We introduce a proposal method that aims to recover segments containing actions in untrimmed videos. Building on techniques for learning sparse dictionaries, we introduce a framework to represent and retrieve activity proposals. We demonstrate the capabilities of our method in not only producing high quality proposals but also in its efficiency.

Contact: fabian.caba@kaust.edu.sa

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

PEOPLE COUNTING BY HUBER LOSS REGRESSION

Cavazza J., Murino V.

Abstract: We address the problem of people counting, proposing a novel regression framework based on a new closed-form solution for the Huber loss. In a multi-view setting, our algorithm adapts itself to the data as to automatically reduce the reconstruction error. We provide a strong theoretical foundation of the method which is able to score superior to several state-of-the-art techniques for people counting on MALL, UCSD and PETS 2016 benchmark datasets.

Contact: jacopo.cavazza@iit.it

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

MOTION ANALYSIS IN DYNAMIC TEXTURES

Chubach O.

Abstract: Dynamic textures contain large amount of details that consume a lot of bitrate when compressed using conventional coding scheme. However, most of those details are irrelevant for people. Therefore, by exploiting human perception properties and manipulating small motions of dynamic textures, one could improve results of compression of such data.

Contact: chubach@ient.rwth-aachen.de

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

ENHANCING EARLY LEARNING IN IMAGE CLASSIFIERS WITH COMPLEX FILTERS

Cotter F., Kingsbury N.

Abstract: We propose to integrate complex, Hilbert symmetric wavelet filters in the early layers of convolutional neural networks (CNNs). Using Hilbert pairs we can have analytic filters, which are more invariant to shifts than their real counterparts. This work is an adaption of the Scatternet developments made by Mallat et al. [1] using Kinsbury's Dual-Tree Complex Wavelet Transform (DTCWT)[2].

Contact: fbc23@cam.ac.uk

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

LEARNING TRANSFERABLE POLICIES FOR AUTONOMOUS MAV CONTROL

Daftry S., Bagnell J. A., Hebert M.

Abstract: The ability to transfer knowledge gained in previous tasks into new contexts is one of the most important mechanisms of human learning. Despite this, adapting autonomous behavior to be reused in partially similar settings is still an open problem in current robotics research. In this work, we propose a generic framework for learning transferable motion policies. We present this in the context of autonomous MAV reactive control.

Contact: shreyanshdaftry219@gmail.com

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

DISCRIMINATIVE CORRELATION FILTERS FOR VISUAL TRACKING

Danelljan M.

Abstract: Discriminative correlation filters (DCF) is a learning technique that efficiently exploits all cyclic shifts of the training samples. Recently DCF methods have shown state-of-the-art results for visual object tracking. This work extends the DCF in several key aspects. First, a spatial regularization is introduced to counter the periodic effects. Second, an online optimization strategy is proposed. Third, a unified learning formulation is introduced for adaptive weighting of the training samples.

Contact: martin.danelljan@liu.se

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

2D IMAGE CLASSIFICATION FOR 3D ANATOMY LOCALIZATION

De Vos B.D., Wolterink J.M., De Jong P.A., Viergever M.A., Išgum I.

Abstract: Localization of anatomical regions of interest (ROIs) is a preprocessing step in many medical image analysis tasks. While trivial for humans, it is complex for automatic methods. In this work convolutional neural networks (CNNs) detect presence of a ROI in 2D orthogonal image slices. The combination of detection results provides a 3D bounding box around the ROI. Results show accurate localization for three ROIs (the heart, aortic arch, and descending aorta) in 100 lung screening CT scans.

Contact: bob@isi.uu.nl

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

AUTOMATIC RECOGNITION OF SURGICAL PHASES

Dergachyova O., Morandi X., Jannin P.

Abstract: Context-aware systems integrated into operating rooms could facilitate clinical routine and help to prevent many of surgery-related adverse events. These systems require an accurate identification of surgical workflow. We propose an automatic fully data-driven and real-time method for detection of surgical phases using video data and instrument usage signals, exploiting no prior knowledge.

Contact: olga.dergachyova@univ-rennes1.fr

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

SOCIAL SIGNAL PROCESSING FOR SHOPPING MALL ROBOT

Deshmukh A., Foster . E., Vinciarelli A.

Abstract: We present a four-year, EU-funded project MuMMER (MultiModal Mall Entertainment Robot) with an overall goal of developing a humanoid robot (based on Aldebaran's Pepper platform) that can interact autonomously and naturally in the dynamic environments of a public shopping mall, providing an engaging and entertaining experience to the general public.

Contact: amol.deshmukh18@gmail.com

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

SUMMARIZING PERSONAL IMAGE COLLECTIONS WITH INTRINSIC PROPERTIES

Erdogan G., Celikkale B., Erdem A., Erdem E.

Abstract: Visual summarization problems are complex intrinsically because definition of a summary has some ambiguity and one correct summary does not exist. A good summary consists of two main properties which are coverage and diversity. Intrinsic properties such as emotions, popularity of images increase in importance depending on prevalence of social media. We proposed an automatic summarization method which considers intrinsic properties of images in addition to coverage and diversity for personal image collection summarization.

Contact: goksuerdogan@cs.hacettepe.edu.tr

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

CONVOLUTIONAL TWO-STREAM NETWORK FUSION FOR VIDEO ACTION RECOGNITION

Feichtenhofer C. Pinz A. Zisserman A.

Abstract: The two-stream architecture [1] incorporates motion information by training separate ConvNets for both appearance in still images and stacks of optical flow. We study a number of ways of fusing ConvNet towers both spatially and temporally in order to best take advantage of this information. We make the following findings: 1. That rather than fusing at the softmax layer, a spatial and temporal network can be fused at a convolution layer without loss of performance, but with a substantial saving in parameters. 2. That it is better to fuse such networks spatially at the last convolutional layer than earlier, and that additionally fusing at the class prediction layer can boost accuracy. 3. That pooling of abstract convolutional features over spatiotemporal neighbourhoods further boosts performance. Based on our studies we propose a new ConvNet architecture for spatiotemporal fusion of video snippets and evaluate its performance on standard benchmarks where it achieves state-of-the-art results.

Contact: feichtenhofer@tugraz.at

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

MULTI-SOURCES BASED PERCEPTION AND LOCALIZATION FOR AUTONOMOUS DRIVING IN URBAN ENVIRONMENTS

Gámez-Serna C., Lombard A., Ruichek Y., Abbas-Turki A.

Abstract: Autonomous vehicle navigation can be broken into two subsystems. The first one referring to a learning phase which includes perceiving and modeling the environment, and localizing the vehicle within the environment. The second one involves planning and deciding the vehicle motion. In this work we will focus on the second subsystem assuming that a planned path already exists. Our method is based on the pure pursuit algorithm considering path's geometry to modify dynamically car's velocity, look-ahead distance, and minimize the lateral error.

Contact: citlalli.gamez-serna@utbm.fr

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

SEMI-INTERACTIVE INCIDENT ANALYSIS - MH17 AS A COMPUTER VISION PROBLEM

Gisolf F., Worring M., Geradts Z.

Abstract: Forensic and accident investigations benefit greatly from photos and videos taken before, during and after events under investigation, However, the analysis of an increasing number of photos and videos takes up a lot of time and effort of the investigator. State-of-the-art computer vision algorithms can help to gain faster and better insight in large multimedia databases. This study applies these algorithms to gain insight in the multimedia database of the MH17 airplane crash investigation (2014). The Dutch Safty Board collected over 60 000 images and over 3 000 videos. THis unique opportunity of analyzing an actual accident investigation database and trying to answer actual questions asked by the investigators allows for the evaluation of the current state-of-the-art. A framework for sensemaking will be developed, using object recognition, identification and detection.

Contact: f.gisolf@safetyboard.nl

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

STEREO VISUAL ODOMETRY BY COMBINING POINTS AND LINES

Gomez-Ojeda, R., Gonzalez-Jimenez, J.

Abstract: Most approaches to stereo visual odometry reconstruct the motion based on the tracking of point features along a sequence of images. However, in low-textured scenes it is often difficult to encounter a large set of point features, or it may happen that they are not well distributed over the image, so that the behavior of these algorithms deteriorates. This paper proposes a probabilistic approach to stereo visual odometry based on the combination of both point and line segment that works robustly in a wide variety of scenarios. The camera motion is recovered through non-linear minimization of the projection errors of both point and line segment features. The method, of course, is computationally more expensive that using only one type of feature, but still can run in real-time on a standard computer and provides interesting advantages, including a straightforward integration into any probabilistic framework commonly employed in mobile robotics.

Contact: rubengooj@gmail.com

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

STRUCTURE TRANSFERRING IMAGE PROCESSING ALGORITHMS

Gracheva I., Kopylov A.

Abstract: An application of Bayesian approach to image processing with structure transferring properties is considered. The framework of Markov random fields makes it possible to take into account the structure, extracted from the "guide" data array, through setting an appropriate probabilistic relationships between the elements of processed data. The proposed approach showed good results in processing quality and computation time by solving the image processing problems such as haze removal problem, HDR image compression and edges refinement of an image.

Contact: gia1509@mail.ru

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

RGBD SEMANTIC SEGMENTATION USING SPATIO-TEMPORAL DATA-DRIVEN POOLING

He Y., Chiu W.-C., Keuper M., Fritz M.

Abstract: In fully convolutional networks, the deconvolutional layer shows different failure modes, like imprecise boundaries or label mistakes. We diagnose these errors as current network aggregate information either too local or too global. Therefore we propose a data-driven pooling layer that integrates with fully convolutional networks and utilizes superpixels. We also extend our method to leverage region-level correspondence across multiple views with a temporal pooling stage.

Contact: yang@mpi-inf.mpg.de

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

OBJECT-BASED LOCALIZATION IN LARGE INDOOR ENVIRONMENTS

Hepp B.

Abstract: Camera localization in repetitive indoor environments such as office buildings is challenging due to lack of distinguishing visual clues and typically changing configurations of objects. We propose a localization approach that relies on the notion of objects and directly incorporates the likelihood of objects being missing, moved around or added to the scene. We investigate possibilities of matching observed objects and their spatial arrangement to a map with a focus on robustness to changing object locations and scalability to large object maps.

Contact: benjamin.hepp@inf.ethz.ch

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

POINT DENSITY-INVARIANT 3D OBJECT DETECTION AND POSE ESTIMATION

Kim S.-A., Yoon K.-J.

Abstract: Among 3D feature descriptors used for describing distinctive and representative features of the object, Point Feature Histograms RGB (PFHRGB) has been evaluated as showing the best performance for 3D object and category recognition. This descriptor, however, is vulnerable to point density variations, which frequently occur in real situations. It produces false correspondences and causes adverse effects on 3D object detection and pose estimation. We tackle this problem and propose an algorithm to find the correct correspondences under the point density variations. We first build multi-scale features on the model point cloud, and then select the right scale by measuring the similarity of the density about each keypoint at the matching stage. In all of the experiments, the proposed algorithm outperforms the original PFHRGB feature descriptor.

Contact: suakim@gist.ac.kr

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

THE WEB-CAM BASED METHOD FOR HEART RATE ESTIMATION

Kopeliovich M., Petrushan M.

Abstract: The remote photoplethysmography method could be used for long-term heart rate monitoring. In this study, the problem of color signal processing is considered. We propose a new signal transformation that generates a signal containing values of the leading coefficients of local quadratic approximation of original signal. The algorithm increases the accuracy of the remote photoplethysmography method by reduction of both high-amplitude fluctuations and high-frequency noise in the signal.

Contact: kop@km.ru

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

SPATIO TEMPORAL MODELLING OF DYNAMIC DEVELOPMENTAL PATTERNS

Licandro R., Langs G., Kampel M.

Abstract: Predictive modelling plays an important role in personalized medicine since it enables the encoding of disease trajectories to optimise an individual's treatment. Aim of the work is to provide a framework, which is capable to learn dynamic developmental trajectories and combines pathological and healthy stadia, to predict, classify or compare them in the image domain. Gaussian Mixture Models and outlier detection techniques are applied to estimate stadia of interest, and longitudinal modelling is performed by designing transition functions to link stadia with each other.

Contact: licandro@caa.tuwien.ac.at

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

INFORMATION SEEKING FOR CRITICAL COMMUNICATIONS

Lopez-Fuentes L., Bagdanov A., van de Weijer J., Skinnemoen H., González-Hidalgo, M.

Abstract: A novel method to optimize bandwidth usage in critical communications is proposed. We develop an active information seeking model that identifies promising regions in low resolution imagery and progressively requests higher resolution regions to perform recognition of higher semantic quality. We develop a car recognition system via identification of its license plate. Results are compared with JPEG. One order of magnitude of bandwidth is saved with little sacrifice in recognition performance.

Contact: laura@ansur.no

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

ORIENTATION DRIVEN BAG OF APPEARANCES FOR PERSON RE-IDENTIFICATION

Ma L., Liu H., Hu L., Wang C., Sun Q.

Abstract: Although person re-identification research is making progress, it still faces many challenges such as varying poses, illumination and viewpoints. For feature representation, existing works usually use low-level descriptors which do not take full advantage of body structure information, resulting in low representation ability. To solve this problem, we propose the mid-level body-structure based feature representation (BSFR) which introduces body structure pyramid for codebook learning and feature pooling in the vertical direction of human body. Besides, varying viewpoints in the horizontal direction of human body usually causes the data missing problem, i.e., the appearances obtained in different orientations of the identical person could vary significantly. To address this problem, the orientation driven bag of appearances (ODBoA) is proposed to utilize person orientation information extracted by orientation estimation technic. To properly evaluate the proposed approach, we introduce two new re-identification datasets, and the experimental results indicate the effectiveness of body structure and orientation information for improving re-identification performance.

Contact: maliqian@sz.pku.edu.cn

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

BLIND EXTRACTION OF FULL-FIELD VIBRATION MODES FROM VIDEO MEASUREMENTS

Yang Y., Dorn C., Mancini T., Talken T., Kenyon G., Farrar C., Mascarenas D.

Abstract: A new framework will be presented for the blind extraction of the full-field, high-resolution, dynamic parameters of an operating structure from video measurements. This high-resolution, full-field dynamic characterization framework opens up a variety of applications that traditionally have not been possible including the ability to accurately localize minute, invisible, structural damage, and a new method enabling realistic video-space, high-fidelity simulations and visualizations/animations of structural dynamics.

Contact: dmascarenas@lanl.gov

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

SPARSE CODING WITH MANIFOLD CONSTRAINTS FOR HUMAN RE-ID

Mirmahboub B., Kiani H., Bhuiyan A., Perina A., Zhang B., Del Bue A., Murino V.

Abstract: Multi-shot images for Human re-identification provide richer information compared to a single-shot image, but produce a high cost of information redundancy. We proposed a novel framework that combines sparse coding and manifold constraints to extract discriminative information from multi-shot images of one pedestrian for person re-identification. The evaluation over two standard multi-shot datasets shows very competitive accuracy of our framework against the state-of-the-art.

Contact: behzad.mirmahboub@iit.it

Presentation Type: Poster

Date: Monday 18 July 2016

Time: 21:30

Poster Session: 1

AUTOMATIC SEGMENTATION OF MR BRAIN IMAGES WITH A CONVOLUTIONAL NEURAL NETWORK

Moeskops P., Viergever M.A., Mendrik A.M., De Vries L.S., Benders M.J.N.L., Išgum I.

Abstract: This work presents an automatic segmentation method for MR brain images using a multi-scale convolutional neural network. It is evaluated on five different data sets: coronal images of preterm infants acquired at 30 and 40 weeks postmenstrual age (PMA), axial images of preterm infants acquired at 40 weeks PMA, axial images of ageing adults acquired at 70 years, and images of young adults acquired at 23 years, resulting in the following average Dice coefficients: 0.87, 0.82, 0.84, 0.86 and 0.91.

Contact: moeskops.pim@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

LIGHT FIELD ANALYSIS FROM UNSTRUCTURED MULTIPLE VIEWS FOR DIMINISHED REALITY

Mori S., Maezawa M., Ienaga N., Tamura H., and Saito H.

Abstract: Diminished reality (DR) is a systematized set of visualization techniques used to hide, delete, or see through undesirable objects in a perceived environment. In this work, we present an object removal technique based on light field analysis from unstructured multiple views for significantly improving the quality and the applicability of DR results. We mainly discuss two DR approaches based on prior and real-time light fields observations.

Contact: s.mori.jp@ieee.org

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

DETECTION OF CHANGES IN FUNCTIONAL HEALTH

Nait Aicha A., Krose B.

Abstract: We present an approach of how to use sensor monitoring to detect changes in the functional health of older adults living alone. By extracting different features from the collected raw sensor data and applying machine learning algorithms, we seek to a continuously monitoring of the functional health.

Contact: a.nait.aicha@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

ROBUST VISUAL SLAM ACROSS SEASONS

Naseer T., Spinello L., Stachniss C., Ruhnke M., Burgard W.

Abstract: Robots require robust lifelong localization, in all visual conditions resilient to seasonal appearances, loops and partially matching trajectories. Our approach achieves all this in a single framework. We extract semi-dense image descriptions and match them for similarities. We formulate a flow network problem and compute matching hypotheses. Furthermore, we use these constraints with robot odometry to formulate a graph-based SLAM problem and compute a jointly maximum likelihood trajectory.

Contact: naseer@cs.uni-freiburg.de

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

UNSUPERVISED LEARNING OF VISUAL REPRESENTATIONS BY SOLVING JIGSAW PUZZLES

Noroozi M., Favaro P.

Abstract: We study the problem of image representation learning without human annotation. We build a convolutional neural network (CNN) that can be trained to solve Jigsaw puzzles as a pretext task, which requires no manual labeling, and then later repurposed to solve object classification and detection. We introduce the context-free network (CFN), a Siamese-ennead CNN. The CFN takes image tiles as input and explicitly limits the receptive field. By training the CFN to solve Jigsaw puzzles, we learn both a feature mapping of object parts as well as their correct spatial arrangement.

Contact: noroozi@inf.unibe.ch

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

ISOMETRIC NON-RIGID SHAPE-FROM-MOTION IN LINEAR TIME

Parashar S., Pizarro D. and Bartoli A.

Abstract: We study Isometric Non-Rigid Shape-from-Motion (Iso-NRSfM): given multiple monocular images, we reconstruct the 3D shape of an object undergoing isometric deformations. We propose a new theoretical framework based on Riemmanian geometry which allows us to derive a system of 2 quartic equations in 2 variables for any number of images. The proposed method outperforms existing work in terms of accuracy and computation cost on synthetic and real datasets.

Contact: shaifali.parashar@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

VISION MEETS TELEDIAGNOSTIC ROBOTICS

Probst T., Fossati A., Van Gool, L.

Abstract: The European FP7 project ReMeDi (grant 610902) aims to develop a system that allows doctors to remotely perform physical and ultrasonography examinations by teleoperating a multifunctional robotic device. In order to perceive the patient and to estimate his/her pain sentiment, the body pose as well as the patient's facial expression is recognized using CV methods. The interactive nature of the application requires the RGB-D data to be processed in real-time. Our research focus lies on developing and evaluating both efficient and robust algorithms to meet these demands.

Contact: probstt@vision.ee.ethz.ch

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

RETINAL IMAGE MOSAICING FOR SLIT-LAMP IMAGING

Prokopetc K.

Abstract: The construction of seamless mosaics from long slit-lamp retinal video sequences is an important and challenging task in navigated Laser Photocoagulation. The main difficulty is accumulated registration drift due to the small number of features away from the optic nerve and the distortion induced by the geometry of the eye and the contact lens. Our work is focused on the development of new techniques for drift-free mosaicing with slit-lamp.

Contact: prokopec.kristina@mail.ru

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

EVALUATION OF WMH SEGMENTATION FOR THE STUDY OF ALZHEIMER'S DISEASE

Rachmadi MF., Valdes Hernandez, M., Komura T.

Abstract: Alzheimer's Disease (AD) is the most common form of dementia, and there is estimated number of 35.6 million people living with dementia according to the 2010 World Alzheimer report. Until now, there is no known treatment for slowing down the progression of this disease because its exact cause has not been discovered yet. We try to recognize lesion or damaged area in the brain as soon as possible in the early stage of AD by segmenting white matter hyperintensities (WMH) in MR images, one of AD biomarkers. However, WMH, sign of AD in MRI, is hard to be segmented especially in the earlier stage because changing of values from normal to damaged area is very subtle.

Contact: febrian.rachmadi@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

CENTERLINE EXTRACTION BY GEODESICS AND GRAPH MODELING

Rahmoun S., Mairesse F., Uji-I H., Hofkens J., Sliwa T.

Abstract: Microscopy imaging techniques often provide corrupted images due to convolution and diffraction effect. The observed polymer chain appears like a thick curve with noise and blur. We propose an automatic method for centerline extraction by geodesics. The shape gives a set of separate geodesics, each one browsing a specific part. To perform the complete curve, we model geodesics spatial relations by graph, and connect them ac-cording to the optimal path traversing the graph

Contact: somia.rahmoun@u-bourgogne.fr

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

EVENT-BASED MULTI-VIEW STEREO

Rebecq H., Gallego, G., Scaramuzza, D.

Abstract: Unlike traditional MVS methods, which address the problem of estimating dense 3D structure from a set of known, sparse viewpoints, EMVS estimates semi-dense 3D structure from a continuously moving event camera with known trajectory. Our EMVS solution elegantly exploits two inherent properties of an event camera: its ability to respond to scene edges, and the fact that it provides continuous measurements as the sensor moves. Our algorithm is able to produce accurate, semi-dense depth maps.

Contact: rebecq@ifi.uzh.ch

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

FASTER R-CNN FEATURES FOR INSTANCE SEARCH

Salvador A.

Abstract: Image representations derived from pre-trained Convolutional Neural Networks (CNNs) have become the new state of the art in computer vision tasks such as instance retrieval. This work explores the suitability for instance retrieval of image- and region-wise representations pooled from an object detection CNN such as Faster R-CNN. We take advantage of the object proposals learned by a Region Proposal Network (RPN) and their associated CNN features to build an instance search pipeline composed of a first filtering stage followed by a spatial reranking. We further investigate the suitability of Faster R-CNN features when the network is fine-tuned for the same objects one wants to retrieve. We assess the performance of our proposed system with the Oxford Buildings 5k, Paris Buildings 6k and a subset of TRECVid Instance Search 2013, achieving competitive results. Code available at: http://imatge-upc.github.io/retrieval-2016-deepvision/

Contact: amaia.salvador@upc.edu

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

MULTISTAGE SFM: A COARSE-TO-FINE APPROACH FOR 3D RECONSTRUCTION

Shah Rajvi, Narayanan P J

Abstract: Several methods have been proposed for large-scale 3D reconstruction from large, unorganized image collections. A large reconstruction problem is typically divided into multiple components which are reconstructed independently using structure from motion (SFM) and later merged together. Incremental SFM methods are popular for basic structure recovery of a single component. They are robust and effective but are mostly sequential in nature. We present a multistage approach for SFM reconstruction of a single component that breaks the sequential nature of the incremental SFM methods. Our approach begins with quickly building a coarse 3D model using only a fraction of features from given images. The coarse model is then enriched by localizing remaining images and matching and triangulating remaining features in subsequent stages. These stages are made efficient and highly parallel by leveraging the geometry of the coarse model. Our method produces similar quality models as compared to other methods while being notably fast and parallel.

Contact: rajvi.a.shah@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

DEEP LEARNING FOR DETECTING MULTIPLE SPACE-TIME ACTION TUBES IN VIDEOS

Singh G., Saha S., Sapienza M., Torr P., Cuzzolin F.

Abstract: In this work we propose a new approach to the spatiotemporal localisation (detection) and classification of multiple concurrent actions within temporally untrimmed videos. Our framework is composed of three stages. In stage 1, a cascade of deep region proposal and detection networks are employed to classify regions of each video frame potentially containing an action of interest. In stage 2, appearance and motion cues are combined by merging the detection boxes and softmax classification scores generated by the two cascades. In stage 3, sequences of detection boxes most likely to be associated with a single action instance, called action tubes, are constructed by solving two optimisation problems via dynamic programming.

Contact: 15056568@brookes.ac.uk

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

IMPLICIT-TO-IMPLICIT REGISTRATION FOR REAL-TIME 3D RECONSTRUCTION FROM RGB-D DATA

Slavcheva M., Kehl W., Navab N., Ilic S.

Abstract: We present a complete pipeline for 3D object reconstruction from RGB-D data, including online frame-to-frame camera tracking, followed by a quick global pose optimization step to ensure optimal geometric consistency. The registration problem in both stages is solved via minimization of a novel implicit-to-implicit energy between pairs of signed distance fields, which avoids expensive correspondence search and is more accurate than ICP-like approaches.

Contact: mira.slavcheva@tum.de

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

REAL-TIME MULTI OBJECT DETECTION AND RECOGNITION USING DEEP LEARNING (CNN)

Sreeramula S., Gan Chun Y.

Abstract: The existing object detection and recognition systems rely on region proposition and classifiers [1] & [2], which are far from real-time requirement. A new real-time multi object detection and recognition (RMODAR) is proposed as a regression problem to spatially isolated bounding boxes and related class probabilities. The proposed R-MODAR processes 60 frames/sec on a GPU, while achieving state-of-the-art object detection accuracy(mAP) on VOC 2007 and 2012 data set are 66.28% mAP and 64.79% mAP.

Contact: sankaraiah.sreeramula@adv-fusionex.com

Presentation Type: Poster

Date: Tuesday 19 July 2001

Time: 21:30

Poster Session: 2

DISCOVERY OF EVERYDAY HUMAN ACTIVITIES FROM LONG-TERM VISUAL BEHAVIOUR USING TOPIC MODELS

Steil J., Bulling A.

Abstract: Human visual behaviour has significant potential for activity recognition and computational behaviour analysis, but previous works focused on supervised methods and recognition of predefined activity classes based on short-term eye movement recordings. We propose a fully unsupervised method to discover users' everyday activities from their long-term visual behaviour combining a bag-of-words representation of visual behaviour that encodes saccades, fixations, and blinks with a LDA topic model.

Contact: jsteil@mpi-inf.mpg.de

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

LEARNING SPATIO-TEMPORAL CO-OCCURRENCE RELATIONSHIP FOR HUMAN ACTION RECOGNITION

Sun Q., Liu M., Liu H.

Abstract: Human action recognition is still a challenging task for computers due to a large variety of complex conditions such as cluttered backgrounds, camera motion and photometric variances of foreground objects. Many human actions, such as "push" and "hug", have similar pose components and hence are easily confused to be one. This problem is called inter-ambiguity. It is more discriminative to model the semantic relationship of local features in terms of normalized google-like distance (NGLD), which measures the co-occurrence (not only neighborhood) frequency of each pair-wise visual words appearing in videos. Human actions make huge senses in moving body parts directionally from one place to another. This reflects the importance of directional information for action representation.

Contact: qsun@mpi-inf.mpg.de

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

USING SPATIAL ORDER TO BOOST THE ELIMINATION OF INCORRECT FEATURE MATCHES

Talker L., Moses Y., Shimshoni I.

Abstract: We propose a method to estimate the number of correct matches between two images without computing them. The set of matches is analyzed using the spatial order of the features, as projected to the x-axis, which reduces the analysis of the matching problem to the analysis of the permutation between the sequences. Using the Kendall distance between the permutations and natural assumptions on the distribution of the correct and the incorrect matches, we estimate the number of correct matches.

Contact: liortalker@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

CONTENT-BASED SPARSITY FOR LOCAL SEMANTIC FEATURES

Tamaazousti Y., Le Borgne H., Popescu A.

Abstract: This work [3] proposes a method to determine automatically a level of sparsity for the semantic features [1, 2] that is adapted to each image content. This method takes into account the amount of information contained by the image through a modeling of the semantic feature entropy and the confidence of the individual values of the feature. We also investigate the use of local regions of the image to constrain them by content-based sparsity and further improve the quality of semantic features.

Contact: youssef.tamaazousti@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

METASTASIS DETECTION IN LYMPH NODE TISSUE SECTIONS USING DEEP LEARNING

Tellez D., Ehteshami B., Litjens G., Karssemeijer N., van der Laak J.

Abstract: In this work, we describe our solution to solve the lesion-based detection task proposed in the ISBI Challenge on cancer metastasis detection in lymph node of 2016. First, a fully convolutional neural network was trained on small patches extracted from whole-slide images (WSI) in the Challenge dataset. Second, this network was used to classify every single pixel of the entire WSIs, and locate individual tumor lesions. Our algorithm yielded remarkable results, achieving the second position in the Challenge leaderboard.

Contact: david.tellezm@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

WEAKLY-SUPERVISED SEMANTIC SEGMENTATION USING MOTION CUES

Tokmakov P., Alahari K., Schmid C.

Abstract: Our work incorporates motion cues in the weakly-supervised motion segmentation framework. This allows to efficiently utilise weakly-annotated videos, achieving the performance on par with the state-of-the-art methods trained on thousands of images with just 150 videos. Training jointly on images and videos archives the new state-of the art performance on the PASCAL VOC 2012 dataset.

Contact: pavel.tokmakov@inria.fr

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

PIXEL-LEVEL ENCODING AND DEPTH LAYERING FOR INSTANCE-LEVEL SEMANTIC LABELING

Uhrig J., Cordts M., Franke U., Brox T.

Abstract: * Our method uses a fully convolutional network (FCN [3]) to densely predict semantic labels, depth and a proposed instance-based encoding * We apply low-level computer vision techniques on the FCN's output for state-of-the-art instance segmentation * Additionally, we predict absolute distances of individual instances from a monocular image * We outperform existing works by a large margin on the street scene datasets KITTI [1] and Cityscapes [2]

Contact: uhrig.jonas@gmail.com

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

HIERARCHICAL 2.5D ALIGNMENT FOR CHANGE DETECTION WITH LARGE VIEWPOINT DIFFERENCES

Van de Wouw D.W.J.M., Dubbelman G., De With P.H.N.

Abstract: Change detection from mobile platforms has many applications, such as countering Improvised Explosive Devices (C-IED). Existing real-time C-IED systems are not robust against large viewpoint differences, which are unavoidable under realistic operating conditions in outdoor environments. To address this, we proposes a hierarchical 2.5D scene-alignment algorithm. This algorithm renders the historic view as if seen by the live camera, combining the accuracy of 2D local image registration with the robustness of 3D scene alignment. It was found that the resulting change detection system detects small changes of only $18 \times 18 \times 9$ cm at distances of 60 meters under large trajectory deviations of up to 2.5 meters.

Contact: d.w.j.m.v.d.wouw@tue.nl

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

SEQUENTIAL MODELS FOR PREDICTIVE DIAGNOSTICS

van der Westhuizen J., Lasenby J.

Abstract: Sequential machine learning techniques allow the extraction of critical information embedded in hours of continuously monitored signals in the ICU. The aim of this project was to provide clinicians with a continuous predictive probability of a patient's physiological status belonging to a range of conditions. Experimentation was done with Hidden Markov Models (HMMs), and Recurrent Neural Networks (RNNs) to determine the best model for prediction, and RNNs showed the most promise to date.

Contact: jv365@cam.ac.uk

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

LONG-TERM TEMPORAL CONVOLUTIONS FOR ACTION RECOGNITION

Varol G., Laptev I., Schmid C.

Abstract: Typical human actions last several seconds and exhibit characteristic spatio-temporal structure. Recent methods attempt to capture this structure and learn action representations with convolutional neural networks. Such representations, however, are typically learned at the level of a few video frames failing to model actions at their full temporal extent. In this work we learn video representations using neural networks with long-term temporal convolutions (LTC). We demonstrate that LTC-CNN models with increased temporal extents improve the accuracy of action recognition. We also study the impact of different low-level representations, such as raw values of video pixels and optical flow vector fields and demonstrate the importance of high-quality optical flow estimation for learning accurate action models. We report state-of-the-art results on two challenging benchmarks for human action recognition UCF101 (92.7%) and HMDB51 (67.2%).

Contact: gul.varol@inria.fr

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

RE-THINKING COMPUTER VISION: EVENT-DRIVEN CORNER DETECTION

Vasco V., Glover A., Bartolozzi C.

Abstract: The detection of consistent feature points in an image is a fundamental task for various kinds of computer vision applications. We present an event-based implementation of the Harris corner detection.

Contact: Valentina.Vasco@iit.it

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

3D OBJECT TACTILE LOCALIZATION AND MODELING

Vezzani G., Pattacini U., Battistelli G., Chisci L., Natale L.

Abstract: Our work addresses 6-DOF tactile localization, i.e. object pose estimation given tactile measurements. We propose a noveland effective Bayesian algorithm, named Memory Unscented Particle Filter (MUPF). The simulated and experimental tests carried out on the iCub humanoid robot show that the algorithm provides accurate and reliable localization and is compatible with real-time requirements.

Contact: giulia.vezzani@iit.it

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

DETECTING HUMAN ACTIVITY EXPLOITING THE REGULARITIES OF BIOLOGICAL MOTION

Vignolo A., Noceti N., Rea F., Sciutti A., Odone F., Sandini G.

Abstract: The capability of detecting human activity in the scene is fundamental for robots which have to collaborate with humans. When the scene is complex and the body part (e.g. the hand) is occluded (e.g. by the used tool), methods which rely on appearance or shape features are likely to fail. To overcome these limitations, we suggest a biologically-inspired human detection method based on motion. We propose a temporal multi-resolution motion description able to deal with the different dynamics of human actions, built on top of low level features that capture biological motion regularities.

Contact: alessia.vignolo@iit.it

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

PREDICTING WHEN SALIENCY MAPS ARE ACCURATE AND EYE FIXATIONS CONSISTENT

Volokitin A., Gygli M., Boix X.

Abstract: Motivation: Understanding which parts of images humans judge to be important helps machines process them. Problem: Saliency maps do not fully characterize viewing patterns. They do not 1) describe whether people have a consensus in viewing images 2) have a confidence measure Research Question Can eye fixation consistency and saliency map accuracy be predicted directly from an image? Result: Both consistency and accuracy can be predicted using image features that encode semantic content

Contact: voanna@vision.ee.ethz.ch

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2

AUTOMATIC CORONARY CALCIUM SCORING IN CARDIAC CT ANGIOGRAPHY USING CONVOLUTIONAL NEURAL NETWORKS

Wolterink J.M., Leiner T., Viergever M.A., Isgum I.

Abstract: We present an algorithm to automatically identify and quantify coronary artery calcification (CAC) in cardiac CT angiography (CCTA). A convolutional neural network identifies potential CAC voxels, which are merged into lesions. These lesions are described by features and classified with a Random Forest. The results show that CAC can be automatically scored in CCTA. A non-contrast CAC scoring CT scan with additional radiation dose, often made prior to CCTA, may be unnecessary.

Contact: j.m.wolterink@umcutrecht.nl

Presentation Type: Poster

Date: Tuesday 19 July 2016

Time: 21:30

Poster Session: 2