

POSTER SESSION BOOKLET



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

University of Catania - University of Cambridge

International Computer Vision Summer School 2017

From Representation to Action and Interaction

Sicily, 9 - 15 July 2017

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:

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

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 2017.

Best Presentation Prize A subset of the submitted posters will be selected by the school committee for short oral presentation. A best presentation prize will be given to the best presentations selected by the school committee.

Scholarship A scholarship 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 2017

*Roberto Cipolla
Sebastiano Battiato
Giovanni Maria Farinella*

List of Posters ¹

1. CELL TRACKING VIA PROPOSAL GENERATION AND SELECTION
Akram S. U., Kannala J., Eklund L., and Heikkilä J.
2. ABNORMAL EVENTS ANALYSIS BY THE LATENT DIRICHLET ALLOCATION (LDA) Dieudonne Fabrice ATREVI, Damien VIVET, Bruno EMILE
3. ORIENTATION REGRESSION OF HAND STRUCTURE IN CLINICAL X-RAY IMAGES BY PRE-TRAINED RESIDUAL NEURAL NETWORKS
Baltruschat I. M., Saalbach A., Nickisch H., Jockel S., Heinrich M. P.
4. FAST IMAGE FILTERING WITH L1 GAUSS TRANSFORM Dina Bashkirova, Shin Yoshizawa, Hideo Yokota
5. LIGHTFIELDS FROM MEGASTEREO Bertel Tobias
6. FFTLASSO: LARGE-SCALE LASSO IN THE FOURIER DOMAIN Adel Bibi, Hani Itani, Bernard Ghanem
7. DRIVER ATTENTION MONITORING THROUGH POSE ESTIMATION
Borghi G., Vezzani R., Cucchiara R.
8. KERNEL MATCHING: CONTINUOUS SHAPE MATCHING WITHOUT ISOMETRY ASSUMPTIONS Vestner M., Lähner Z., Boyarski A., Slossberg R., Remez T., Litany O., Rodola E., Bronstein M.A., Bronstein M.M, Cremers D., Kimmel R.
9. X-RAY POSENET: 6 DOF POSE ESTIMATION FOR MOBILE X-RAY DEVICES Bui M., Albarqouni S., Schrapp M., Navab N., Ilic S.
10. SE3-NETS: LEARNING RIGID BODY MOTION USING DEEP NEURAL NETWORKS Byravan A., Fox D.

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

-
11. ONE-SHOT VIDEO OBJECT SEGMENTATION Caelles S., Maninis K., Pont-Tuset J., Leal-Taixé L., Cremers D., Van Gool L.
 12. DEEP NEURAL NETWORKS UNDER STRESS Carvalho M., Cord M., Avila S., Thome N., Valle E.
 13. OPTICAL FLOW BASED DEPTH ESTIMATION FROM LIGHT FIELD Chen Y., Alain M., Smolic A.
 14. PREDICTING HUMAN EYE FIXATIONS VIA AN LSTM-BASED SALIENCY ATTENTIVE MODEL Cornia M., Baraldi L., Serra G., Cucchiara R.
 15. LOCAL SHAPE SPECTRUM ANALYSIS FOR 3D FACIAL EXPRESSION RECOGNITION Derkach D., Sukno F.
 16. LAPLACIAN PYRAMID OF VARIATIONAL AUTOENCODERS Dorta G., Vicente S., Agapito L., Campbell N., Prince S., Simpson I.
 17. WHO'S BETTER, WHO'S BEST: SKILL DETERMINATION IN VIDEO USING DEEP RANKING Doughty H., Damen D., Mayol-Cuevas W.
 18. DEEP SINGLE AND DIRECT MULTI-VIEW DEPTH FUSION Fácil, J. M., Concha, A., Montesano, L., Civera, J.
 19. ISCHEMIC STROKE ENHANCEMENT USING A VARIATIONAL MODEL AND THE EXPECTATION MAXIMIZATION METHOD Alves, A; Jenane, R; Pina, D.
 20. INVESTIGATING NUISANCE FACTORS IN FACE RECOGNITION WITH DCNN REPRESENTATION Ferrari C., Lisanti G., Berretti S., Del Bimbo A.
 21. INTERPRETABLE EXPLANATIONS OF BLACK BOXES BY MEANINGFUL PERTURBATION Fong R., Vedaldi, A.
 22. LEARNING VISUAL SIMILARITY FOR IMAGE RETRIEVAL Garcia N., Vogiatzis G.

-
23. AUTOMATED MEDICAL IMAGE CAPTION GENERATION Gasimova, A., Rueckert, D.
 24. SPACE-TIME PRIMITIVES TO UNDERSTAND HUMAN MOVEMENTS Goyal G, Malafronte D, Vignolo A, Noceti N, Odone F, DeVito E
 25. DEEP LEARNING FOR SENSOR FUSION IN ADVERSE WEATHER CONDITIONS Gruber T., Bijelic M., Ritter W., Ditzel C., Pfeuffer A., Dietmayer K.
 26. CV AND ROBOTICS FOR HUMAN CENTERED SYSTEMS Gulde T., Browatzki B., Ludl D., Curio C.
 27. TINY HEAD POSE CLASSIFICATION BY BODILY CUES Hasan I., Tsesmelis T., Galasso F., Bue D. A., and Cristani M.
 28. SHOT BOUNDARY DETECTION THROUGH SPATIO-TEMPORAL CNNs Ahmed M. Hassaniien, Mohamed Elgharib, Ahmed Selim, Mohamed Hefeeda, Wojciech Matusik
 29. RATE-ADAPTIVE GLOBAL IMAGE REPRESENTATIONS Heisterklaus I., Ereemeev N., Bagana R.
 30. MOTION TRACKING OF CELLULAR OBJECTS: USING HOUGH, MSER, SNAKES, LM SOLVER, KALMAN FILTERING Kapoor Varun, Preibisch Stephan, Reber Simone
 31. DETECTION AND DEBLURRING OF FAST MOVING OBJECTS IN VIDEOS Kotera J., Bilkova Z., Matas J., Rozumnyi D.
 32. MODELING ARTICULATIONS FOR HUMAN SHAPE RECOVERY Kravchenko E., Galun M., Basri R.
 33. DEEP INCREMENTAL SCENE UNDERSTANDING Laina I., Rupprecht C., Tateno K., Belagiannis V., Tombari F., Navab N.
 34. CLASS-ADAPTED BLIND IMAGE DEBLURRING Ljubenovic M., Figueiredo M.

-
35. TOWARDS PEDESTRIAN INTENTION ESTIMATION Ludl D., Browatzki B., Gulde T., Curio C.
 36. A LOCATION-AWARE VLAD FOR LANDMARK RECOGNITION Magliani F.
 37. LEARNING SEGMENTATION OF DEPTH MAPS FROM 3D CAD MODELS Mahé H., Marraud D., Comport A. I.
 38. DEEP CNNs FOR OBJECT DETECTION ON THE ICUB ROBOT Maittini E., Pasquale G., Rosasco L., Natale L.
 39. VIDEO SIGNAL ANALYSIS BASED ON 3D SHEARLETS Malafrente D., Odone F., De Vito E.
 40. LEARNING DEEP NBNN REPRESENTATIONS FOR ROBUST PLACE CATEGORIZATION Mancini M., Rota Bulò S., Ricci E., Caputo B.
 41. SINGLE-IMAGE SUPER RESOLUTION FOR STREAMING VIDEO VIA CONVOLUTIONAL NEURAL NETWORK Morgacheva A., Kulikov V.
 42. PERSISTENT AERIAL TRACKING FROM UAVS Mueller M., Smith N., Ghanem B.
 43. FROM BENEDICT CUMBERBATCH TO SHERLOCK HOLMES - CHARACTER IDENTIFICATION IN TV SERIES Nagrani A., Zisserman A.
 44. TOWARDS A VISUAL PRIVACY ADVISOR: UNDERSTANDING AND PREDICTING RISKS IN IMAGES Orekondy T., Schiele B., Fritz M.
 45. A UNIFIED APPROACH OF MULTI-SCALE DEEP AND HAND-CRAFTED FEATURES FOR DEFOCUS ESTIMATION Park Jinsun, Tai Yu-Wing, Cho Donghyeon, Kweon In So
 46. VIRALIENCY: POOLING LOCAL VIRALITY Pilzer A., Alameda-Pineda X., Xu D., Sebe N., Ricci E.

-
47. MULTIVIEW MACHINE LEARNING USING AN ATLAS OF CARDIAC CYCLE MOTION Puyol Esther, Langet Helene, Schnabel Julia, De Craene Mathieu, Piro Paolo, King Andrew P.
 48. USING LSTM TO REDUCE COMPUTATIONAL COST IN DEEP CONVOLUTIONAL NEURAL NETWORK Ray A., Sakaue F., Sato J.
 49. INTEGRATING STEREO VISION WITH A CNN TRACKER FOR A PERSON-FOLLOWING ROBOT Sahdev R., Chen B.X., Tsotsos J.K.
 50. ENHANCING GRASSMANN DISCRIMINANT ANALYSIS Souza L., Gatto B., Hino H., Fukui K.
 51. ONLINE SELF-SUPERVISED FCN FOR FREE-SPACE DETECTION WITH INTELLIGENT VEHICLES Sanberg W., Dubbelman G., de With, P.
 52. SURFACE INSPECTION OF OPTICAL ELEMENTS Schöch A., Perez P., Linz-Dittrich S., Bach C., Ziolk C.
 53. DEEPWARP: PHOTOREALISTIC IMAGE RESYNTHESIS FOR GAZE MANIPULATION Ganin Y., Kononenko D., Sungatullina D., Lempitsky V.
 54. MATCHMOVING ON SET WITH THE USE OF REAL-TIME VISUAL-INERTIAL LOCALIZATION AND DEPTH ESTIMATION Tarko J., Campbell N.
 55. TWO-STREAM CONVOLUTIONAL NETWORKS FOR DYNAMIC TEXTURE SYNTHESIS Tesfaldet Matthew., Brubaker Marcus A., Derpanis Konstantinos G.
 56. BEYOND BACTERIA ANALYSIS WITH DEEP LEARNING AND HYPERSPECTRAL DATA Turra G., Arrigoni S., Signoroni A.
 57. ONLINE ADAPTATION OF CONVOLUTIONAL NEURAL NETWORKS FOR VIDEO OBJECT SEGMENTATION Voigtlaender P., Leibe B.

-
58. ROBUST REGISTRATION FRAMEWORK OF MULTI-MODALITY MR DATA Wang C, MacGillivray T, Koutraki Y, Boardman J, Sparrow S, Pataky R, and Semple S
 59. SILNET: 3D SHAPE FROM SILHOUETTES Wiles O., Zisserman A.
 60. WHAT ARE YOU LOOKING AT IN A MUSEUM Yang J., Sörös G.

CELL TRACKING VIA PROPOSAL GENERATION AND SELECTION

Akram S. U., Kannala J., Eklund L., and Heikkilä J.

Abstract: Cell tracking plays a vital role in understanding biological processes involved in development and disease by enabling analysis of (morpho)dynamic cell behavior. Current methods are heavily engineered for specific applications and lack generalization. We present a general cell tracking method which: 1) Generates cell proposals using CNNs. 2) Links them using edges representing cellular events. 3) Selects globally optimal subset of cell proposals & their associations, providing cell tracks.

Contact: saad.akram@oulu.fi

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

ABNORMAL EVENTS ANALYSIS BY THE LATENT DIRICHLET ALLOCATION (LDA)

Dieudonne Fabrice ATREVI, Damien VIVET, Bruno EMILE

Abstract: In this work, we investigated the abnormal events detection in video using visual saliency for Spatio-temporal interest point selection, combined with the Latent Dirichlet Allocation for event modeling. The point-of-interest filtering is carried out using the saliency score, allowing only those with visual importance to be considered. A model of normal events is learned thanks to the probabilistic generative model "Latent Dirichlet Allocation" (LDA), known for its performance in textual data mining. The detection of an abnormal or rare event is carried out in a probabilistic way via the learned model. The methodology has been applied successfully on the USCD Ped1 and UMN datasets and results are compared to the state of the art.

Contact: fabrice.atrevi@univ-orleans.fr

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

ORIENTATION REGRESSION OF HAND STRUCTURE IN CLINICAL X-RAY IMAGES BY PRE-TRAINED RESIDUAL NEURAL NETWORKS

Baltruschat I. M., Saalbach A., Nickisch H., Jockel S., Heinrich M. P.

Abstract: An upright orientation of the anatomy in a digital X-ray image is preferred for consistency and quality reasons by the majority of radiologists. In almost half of the clinical cases, the anatomy is not upright orientated, which is why the images have to be digitally rotated by radiographers. Earlier work has shown that automated orientation detection resulted in low error rates, but required specially designed algorithms for individual anatomies. In this work, we propose a new approach to overcome time-consuming feature engineering by means of Residual Neural Networks (ResNet), which extract generic low-level and high-level features, and provide promising options for medical imaging. Our method uses the extracted features to estimate the orientation via linear regression, and can be further improved by fine-tuning the ResNet layers. The method was evaluated on 926 hand X-ray images and achieves a state-of-the-art mean absolute error of 2.79%

Contact: i.baltruschat@uke.de

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

FAST IMAGE FILTERING WITH L1 GAUSS TRANSFORM

Dina Bashkirova, Shin Yoshizawa, Hideo Yokota

Abstract: Gaussian convolution is a core tool in many applications in computer vision and image processing. Since computing the exact Gaussian convolution on large datasets is a highly expensive task, a fast and accurate approximation technique is needed. We propose a novel 2D approximation approach based on L1 distance metric and domain splitting that allows to significantly speed up the computational process compared to conventional methods while preserving high accuracy.

Contact: dina.bashkirova@riken.jp

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

LIGHTFIELDS FROM MEGASTEREO

Bertel Tobias

Abstract: This poster describes a method to augment restricted lightfield functionality on-the-fly exploiting a ray database created for cylindrical stereoscopic panoramas.

Cylindrical panoramas capture an environment in one single image. Lightfields allow changing the viewing direction and focus of a captured scene.

Contact: T.B.Bertel@bath.ac.uk

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

FFTLASSO: LARGE-SCALE LASSO IN THE FOURIER DOMAIN

Adel Bibi, Hani Itani, Bernard Ghanem

Abstract: The LASSO problem has been studied and used in a variety of areas, ranging from signal processing and information theory to computer vision and machine learning. In the vision community, it found its way into applications, including face recognition, tracking, super resolution, to name a few. This paper proposes a novel circulant reformulation of the LASSO that lifts the problem to a higher dimension, where ADMM can be efficiently applied to its dual form. Because of this lifting, all optimization variables are updated using only basic element-wise operations, the most computationally expensive of which is a 1D FFT. In this way, there is no need for a linear system solver nor matrix-vector multiplication. Since all operations in our FFTLasso method are element-wise, the subproblems are completely independent and can be trivially parallelized (e.g. on a GPU). The attractive computational properties of FFTLasso are verified by extensive experiments on synthetic and real data and on the face recognition task.

Contact: adel.bibi@kaust.edu.sa

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

DRIVER ATTENTION MONITORING THROUGH POSE ESTIMATION

Borghi G., Vezzani R., Cucchiara R.

Abstract: Driver inattention is one of the most important factor in road crashes [1]. Defining driver attention is not easy since it is classified in different typologies, confused with fatigue [3] and is strictly connected with final real world applications developed by car companies. We develop a framework for accurate and real time driver pose estimation, as indicator of attention level. We collect a new dataset to test the system. Results show that our method overcomes all recent state-of-art works.

Contact: guido.borghi@unimore.it

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

KERNEL MATCHING: CONTINUOUS SHAPE MATCHING WITHOUT ISOMETRY ASSUMP- TIONS

Vestner M., Löhner Z., Boyarski A., Slossberg R., Remez T., Litany O., Rodola E., Bronstein M.A., Bronstein M.M, Cremers D., Kimmel R.

Abstract: We present a method for finding correspondence between two manifolds based on matching their heat kernels. Despite the combinatorial nature of the problem, we provide a simple algorithm with an appealing interpretation.

Contact: amitboy@tx.technion.ac.il

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

X-RAY POSENET: 6 DOF POSE ESTIMATION FOR MOBILE X-RAY DEVICES

Bui M., Albarqouni S., Schrapp M., Navab N., Ilic S.

Abstract: Precise reconstruction of 3D volumes from X-ray projections requires precisely pre-calibrated systems where accurate knowledge of the systems geometric parameters is known. When dealing with mobile X-ray devices such calibration parameters are unknown. In industrial applications, nominal CAD models of the object to be reconstructed are usually available. We rely on this prior information and employ deep learning to learn the mapping between X-ray projections and its pose.

Contact: mai.bui@tum.de

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

SE3-NETS: LEARNING RIGID BODY MOTION USING DEEP NEURAL NETWORKS

Byravan A., Fox D.

Abstract: We introduce the SE3-Net, a deep neural network designed to model rigid body motion from raw point cloud data. Based only on sequences of depth images, action vectors, and point wise data associations, SE3-Nets learn to segment affected object parts and predict their motion, parameterized as SE(3) transforms. We show that SE3-Nets perform significantly better than a traditional flow network on many simulated scenarios and on a real world scenario with a Baxter robot pushing objects on a table.

Contact: barun@cs.washington.edu

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

ONE-SHOT VIDEO OBJECT SEGMENTATION

Caelles S., Maninis K., Pont-Tuset J., Leal-Taixé L., Cremers D., Van Gool L.

Abstract: We provide a fast and accurate CNN architecture for the task of semi-supervised Video Object Segmentation: 1. A fully-convolutional neural network architecture learns the appearance of a single annotated object. 2. Temporal consistent and stable results even though all frames are process independently. 3. Easy to work with any tradeoff between accuracy and execution speed, depending on the application. 4. State-of-the-Art results in the DAVIS and Youtube datasets.

Contact: scaelles@vision.ee.ethz.ch

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

DEEP NEURAL NETWORKS UNDER STRESS

Carvalho M., Cord M., Avila S., Thome N., Valle E.

Abstract: Deep architectures have been used for transfer learning with state-of-the-art performance, but their features remain largely unstudied under the transfer perspective. We present an extensive analysis of their resiliency, focusing on the trade-off between performance and compression. By reducing their precision and dimensionality, we show they are more robust to these disturbances when compared to classical approaches, achieving a compression rate of 98.4%, while losing only 0.88% of their score.

Contact: micael.carvalho@lip6.fr

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

OPTICAL FLOW BASED DEPTH ESTIMATION FROM LIGHT FIELD

Chen Y., Alain M., Smolic A.

Abstract: A novel depth estimation method from light field is proposed based on an advanced feature flow estimation method. This estimator is applied on each sequence of images taken along an angular dimension of the light field, which produces several corresponding sequence of disparity map estimates. Thanks to edge-aware filtering in feature flow, the disparity map estimates are highly consistent, which allows a fast and simple aggregation step to create a single final depth map. We can also aggregate all depth maps sequence to create one extra dense point cloud in 3D space.

Contact: tripleYang.Chen@gmail.com

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

PREDICTING HUMAN EYE FIXATIONS VIA AN LSTM-BASED SALIENCY ATTENTIVE MODEL

Cornia M., Baraldi L., Serra G., Cucchiara R.

Abstract: In this work we go beyond standard approaches to saliency prediction and we present a novel model which can predict accurate saliency maps by incorporating neural attentive mechanisms. The core of our solution is a Convolutional LSTM that focuses on the most salient regions of the input image to iteratively refine the predicted saliency map. Additionally, to tackle the center bias present in human eye fixations, our model can learn a set of prior maps generated with Gaussian functions.

Contact: marcella.cornia@unimore.it

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

LOCAL SHAPE SPECTRUM ANALYSIS FOR 3D FACIAL EXPRESSION RECOGNITION

Derkach D., Sukno F.

Abstract: We investigate the problem of facial expression recognition using 3D data. Building from one of the most successful frameworks for facial analysis using exclusively 3D geometry, we extend the analysis from a curve-based representation into a spectral representation. Spectral representations are based on the decomposition of the geometry in its spatial frequency components, much like a Fourier transform, which are related to intrinsic characteristics of the surface. In this work, we propose the use of spectrum based on Graph Laplacian operator. Our results confirm that the proposed approach produces consistently higher recognition rates than the curves-based approach, thanks to a more complete description of the surface, while requiring a lower computational complexity.

Contact: dmytro.derkach@upf.edu

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

LAPLACIAN PYRAMID OF VARIATIONAL AUTOENCODERS

Dorta G., Vicente S., Agapito L., Campbell N., Prince S., Simpson I.

Abstract: We present LapCVAE, a new generative model, which decomposes the image generation process into smaller tractable steps. Our method achieves lower reconstruction error than the original VAE [1] model, and allows for novel image editing applications that take advantage of the coarse to fine structure of the model.

Contact: G.Dorta.Perez@bath.ac.uk

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

WHO'S BETTER, WHO'S BEST: SKILL DETERMINATION IN VIDEO USING DEEP RANKING

Doughty H., Damen D., Mayol-Cuevas W.

Abstract: We present a method for assessing skill of performance from video, for a variety of tasks ranging from drawing to surgery. We formulate the problem as pairwise and overall ranking of video collections, and propose a supervised deep ranking model to learn discriminative features between pairs of videos exhibiting different amounts of skill. We utilise a two-stream Temporal Segment Network to capture both the type and quality of motions and the evolving task state.

Contact: hazel.doughty@bristol.ac.uk

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

DEEP SINGLE AND DIRECT MULTI-VIEW DEPTH FUSION

Fácil, J. M., Concha, A., Montesano, L., Civera, J.

Abstract: Dense and accurate 3D mapping from a monocular sequence is a key technology for several applications and still an open research area. This work leverages recent results on single-view CNN-based depth prediction and fuses them with multi-view depth estimation. Both approaches present complementary strengths. Multi-view depth is highly accurate but only in high-texture areas and high-parallax cases. Single-view depth captures the local structure of mid-level regions, including texture-less areas, but the estimated depth lacks global coherence. The single and multi-view fusion we propose is challenging in several aspects. First, both depths are related by a deformation that depends on the image content. Second, the selection of multi-view points of high accuracy might be difficult for low-parallax configurations. We present contributions for both problems. Our results in the public datasets of NYUv2 and TUM shows that our algorithm outperforms the individual single and multi-view approaches.

Contact: jmfacil@unizar.es

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

ISCHEMIC STROKE ENHANCEMENT USING A VARIATIONAL MODEL AND THE EXPECTATION MAXIMIZATION METHOD

Alves, A; Jennane, R; Pina, D.

Abstract: Correctly diagnosing the extension of stroke requires a high efficiency to avoid irreversible cerebral damage. This work presents a novel approach to enhance the visual perception of acute ischemic stroke in Non-Enhanced Computed Tomography (NECT). A computational algorithm was proposed using different image processing techniques, such as a variational model decomposition and the expectation maximization method. Our main objective was to enable less experienced physicians to reliably detect early signs of acute ischemic stroke.

Contact: allan_ffa@hotmail.com

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

INVESTIGATING NUISANCE FACTORS IN FACE RECOGNITION WITH DCNN REPRESENTATION

Ferrari C., Lisanti G., Berretti S., Del Bimbo A.

Abstract: Deep learning based approaches proved to be effective in many computer vision applications, including 'face recognition in the wild'. The performance of a DCNN are surely influenced by the characteristics of the data used to train the network. In this work, we evaluate the effect of the bounding box size, alignment, positioning and data source on face recognition with DCNNs.

Contact: ferrari.claudio88@gmail.com

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

INTERPRETABLE EXPLANATIONS OF BLACK BOXES BY MEANINGFUL PERTURBATION

Fong R., Vedaldi, A.

Abstract: Recently developed image saliency methods that explain where CNNs “look” in an image for a given task are limited by their heuristic nature and architectural constraints. In this work, we make two main contributions: First, we propose a general framework for learning different kinds of explanations for any black box algorithm. Second, we introduce a paradigm that learns the minimally salient part of an input by directly editing it and learning from the affected changes to its output.

Contact: ruthfong@robots.ox.ac.uk

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

LEARNING VISUAL SIMILARITY FOR IMAGE RETRIEVAL

Garcia N., Vogiatzis G.

Abstract: Can a computer vision system learn the concept of visual similarity? Traditionally, visual similarity in image retrieval is measured by using standard metrics, such as Euclidean distance or cosine similarity. However, these metrics are independent from data and might be missing the nonlinear structure of visual representations. In this work, we propose to learn a nonlinear visual similarity function directly from image representations by optimizing a neural network model.

Contact: garciadn@aston.ac.uk

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

AUTOMATED MEDICAL IMAGE CAPTION GENERATION

Gasimova, A., Rueckert, D.

Abstract: Creating suitable training data and ground truths for the purpose of machine learning is a challenging and time consuming task in the medical domain. There is potential for a learning framework that can attempt to leverage the vast amount of clinical reports made by radiologists, together with their corresponding images, to detect diseases and their contexts (e.g. severity, location) based on recent progress in image captioning for computer vision applications. We propose to take inspiration from the recent success of saliency models in order to perform automated image caption generation and object recognition in fetal Ultrasound.

Contact: a.gasimova16@imperial.ac.uk

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

SPACE-TIME PRIMITIVES TO UNDERSTAND HUMAN MOVEMENTS

Goyal G, Malafronte D, Vignolo A, Noceti N, Odone F, DeVito E

Abstract: This work investigates the application of Shearlet Transform, with the use of appropriately learned space time primitives for modelling upper body actions observed from multiple views.

Contact: gaurvigoyal@gmail.com

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

DEEP LEARNING FOR SENSOR FUSION IN ADVERSE WEATHER CONDITIONS

Gruber T., Bijelic M., Ritter W., Ditzel C., Pfeuffer A., Dietmayer K.

Abstract: Current driver assistance systems often malfunction or even fail in adverse weather where the accident risks are highest. The EU project DENSE focusses on developing and demonstrating an all-weather sensor suite for traffic services, driver assistance and autonomous driving. In addition to innovative sensor technologies, the signals of different sensors (radar, LIDAR, camera) are fused by convolutional neural networks (CNNs) on a very low sensor signal level in order to enable reliable environment perception in any weather condition.

Contact: tobias.gruber@daimler.com

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

CV AND ROBOTICS FOR HUMAN CENTERED SYSTEMS

Gulde T., Browatzki B., Ludl D., Curio C.

Abstract: Computer vision and robotics are both versatile fields of research and have received a lot of attention for the past decades. By combining well-established methods like object recognition, motion- and body-tracking or robotics-navigation and -perception we are now able to face daily tasks in a new and human-centered way. By employing alternative interface technologies, such as vibro-tactile feedback, we're even able to enlarge the conventional application domain and support people with special needs.

Contact: thomas.gulde@reutlingen-university.de

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

TINY HEAD POSE CLASSIFICATION BY BODILY CUES

Hasan I., Tsesmelis T., Galasso F., Bue D. A., and Cristani M.

Abstract: Head pose is an important cue for computer vision. Traditionally considered in human computer interaction applications, it becomes very hard to model in surveillance scenarios, due to the tiny head size. Here we present a framework based on Faster RCNN [1], which introduces a novel branch for head pose estimation. The key idea is to leverage the presence of people body, to better infer the head pose through a joint optimization process. Additionally in this work, we present that the head pose captures people focus also known as Visual Frustum of Attention (VFOA). We illustrate how VFOA could be use for generating attention map of the scene, which provides a better scene understanding

Contact: irtiza.hasan@univr.it

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

SHOT BOUNDARY DETECTION THROUGH SPATIO-TEMPORAL CNNs

Ahmed M. Hassanien, Mohamed Elgharib, Ahmed Selim, Mohamed Hefeeda, Wojciech Matusik

Abstract: Shot boundary detection (SBD) is an important preprocessing step for video manipulation. Here, each segment of frames is classified as either sharp, gradual or no transition. Current SBD techniques analyze hand-crafted features and attempt to optimize both detection accuracy and processing speed. However, the heavy computations of optical flow prevents this. To achieve this aim, we present an SBD technique based on spatio-temporal Convolutional Neural Networks (CNN). Since current datasets are not large enough to train an accurate SBD CNN, we present a new dataset containing more than 3.5 million frames of sharp and gradual transitions. The transitions are generated synthetically using image compositing models. Our dataset contain additional 70,000 frames of important hardnegative no transitions. We perform the largest evaluation to date for one SBD algorithm, on real and synthetic data, containing more than 4.7 million frames. In comparison to the state of the art, we outperform dissolve gradual detection, generate competitive performance for sharp detections and produce significant improvement in wipes. In addition, we are up to 11 times faster than the state of the art.

Contact: ahassanien@hbku.edu.qa

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

RATE-ADAPTIVE GLOBAL IMAGE REPRESENTATIONS

Heisterklaus I., Ereemeev N., Bagana R.

Abstract: In Mobile Visual Search Applications, a compact image representation which can adapt to changes in the network is of utmost importance. One possible global descriptor is the Fisher Vector. We show that an adaptive Fisher Vector can be developed by using an adaptive Gaussian Mixture Model. Different strategies for building an adaptive model are currently under review. We present first results for a simple pruning strategy. They are promising but open questions remain.

Contact: heisterklaus@ient.rwth-aachen.de

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

MOTION TRACKING OF CELLULAR OBJECTS: USING HOUGH, MSER, SNAKES, LM SOLVER, KALMAN FILTERING

Kapoor Varun, Preibisch Stephan, Reber Simone

Abstract: Design a tool to precisely track microtubule (MT) dynamics in TIRF microscopy. Cells/ Nuclei tracking to obtain Intensity fluctuations/ shape factor changes over time. Live demo of all the tools available on the Laptop.

Contact: varun.kapoor@mdc-berlin.de

Presentation Type: Poster

Date: Monday 10 July 2017

Time: 21:30

Poster Session: 1

DETECTION AND DEBLURRING OF FAST MOVING OBJECTS IN VIDEOS

Kotera J., Bilkova Z., Matas J., Rozumnyi D.

Abstract: Objects that move at high velocities appear in photos and videos only as blurred streaks, much larger than the object's size. We call these the fast moving objects (FMOs). Detection and tracking of such objects is difficult, because the object's appearance is so blurred that traditional methods typically fail. We present of method for detection and subsequent tracking of such objects in video sequences. On top of that, we present a method for recovery of the object's appearance (deblurring).

Contact: kotera@utia.cas.cz

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

MODELING ARTICULATIONS FOR HUMAN SHAPE RECOVERY

Kravchenko E., Galun M., Basri R.

Abstract: We introduce a method for 3D human pose estimation and shape recovery from a pair of aligned depth and color images. Parameterizing articulation in linear blend skinning model by bone transformation matrices, we constrain them to be near similarity transformations by controlling its singular values. Color information is used to obtain accurate segmentation [2] and 2D joint detection [3]. We present preliminary results on Berkeley MHAD [1] dataset.

Contact: ekaterina.kravchenko@weizmann.ac.il

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

DEEP INCREMENTAL SCENE UNDERSTANDING

Laina I., Rupprecht C., Tateno K., Belagiannis V., Tombari F., Navab N.

Abstract: This poster shows a composite of two methods combined to perform incremental scene understanding with a single camera. We present a CNN architecture for real-time monocular depth estimation that can then be deployed for accurate and dense monocular reconstruction. The dense depth maps are fused together with depth data from direct monocular SLAM and also with deep learned semantic labels, yielding a semantically coherent scene reconstruction with absolute scale from a single camera.

Contact: iro.laina@tum.de

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

CLASS-ADAPTED BLIND IMAGE DEBLURRING

Ljubenovic M., Figueiredo M.

Abstract: Most of blind image deblurring methods are focused on natural images. Specific classes of images (e.g., text, face, fingerprints) are found in many important applications, but methods tailored for natural images do not take into consideration the particular structure of these images. We proposed a method with a patch-based class-adapted image prior suitable for blind deblurring of images that do belong to specific classes.

Contact: mlju@lx.it.pt

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

TOWARDS PEDESTRIAN INTENTION ESTIMATION

Ludl D., Browatzki B., Gulde T., Curio C.

Abstract: Recognizing human actions and reliably inferring their meaning is one of the core challenges in machine vision. Reliable perception processes for automated urban environment sensing are still in desperate need. Most data-driven algorithms are trained on standard datasets. Typical urban human activities are underrepresented in such datasets despite the fact, that annotation is difficult to obtain. We present a new simulation framework dedicated to produce realistic sensor data for human movements in various contexts allowing to investigate and improve existing algorithms.

Contact: Dennis.Ludl@reutlingen-university.de

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

A LOCATION-AWARE VLAD FOR LANDMARK RECOGNITION

Magliani F.

Abstract: Landmark recognition is a challenging task and it is still a very active field of research. Changes in the image resolution, illumination conditions, viewpoint and the presence of distractors such as trees or traffic signs (just to mention some) make the task of matching features between a query image and the database rather difficult.

Contact: federico.magliani@studenti.unipr.it

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

LEARNING SEGMENTATION OF DEPTH MAPS FROM 3D CAD MODELS

Mahé H., Marraud D., Comport A. I.

Abstract: In the context of a robot evolving in an aircraft assembly line, this poster introduces the semantic segmentation of depth maps where, instead of training a deep neural networks with a mass of real labelled data, we generate synthetic training samples exploiting a pre-existing virtual model of the scene.

Contact: hmahe.pro@gmail.com

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

DEEP CNNs FOR OBJECT DETECTION ON THE ICUB ROBOT

Maiettini E., Pasquale G., Rosasco L., Natale L.

Abstract: In this work we explore the application of latest Deep Learning (DL) methods for object detection to provide a humanoid robot, like iCub[3], with the ability to learn to localize and recognize objects in the real world. While DL recently reported great results on large-scale datasets, applying these methods to a robotic setting comes with some challenges: firstly, the lack of manual annotations, but also variations in light/viewpoint/background, caused by the dynamic environment around the robot. This latter, on the other hand, is a great deal of contextual information. In this regard, we show how it is possible to use the depth of the scene and human interaction to collect sufficient ground truth to effectively train a Region-CNN that, deployed on the robot, provides a robust detection system, able to generalize to different settings.

Contact: elisa.maiettini@iit.it

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

VIDEO SIGNAL ANALYSIS BASED ON 3D SHEAR- LETS

Malafrente D., Odone F., De Vito E.

Abstract: In our research we exploit the properties of the 3D Shearlet Transform to address the problem of analysing 2D+T signals. Our work focuses on the detection and description of meaningful spatio-temporal interest points while at the same time estimating the motion occurring in the scene.

Contact: damiano.malafrente@dibris.unige.it

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

LEARNING DEEP NBNN REPRESENTATIONS FOR ROBUST PLACE CATEGORIZATION

Mancini M., Rota Bulò S., Ricci E., Caputo B.

Abstract: We present a model for semantic place categorization using data obtained from RGB cameras. Inspired by the benefits showed by previous part-based models, we propose an architecture which fully integrates local representations, derived from pre-trained Convolutional Neural Networks (CNNs), and a Naïve Bayes Nearest Neighbor (NBNN) model for classification. Experimental results show that our algorithm outperforms previous methods and it is robust to occlusions, environmental and sensor changes.

Contact: mancini@dis.uniroma1.it

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

SINGLE-IMAGE SUPER RESOLUTION FOR STREAMING VIDEO VIA CONVOLUTIONAL NEURAL NETWORK

Morgacheva A., Kulikov V.

Abstract: Recovering a high-resolution image or a video from a low-resolution source is a subject of increased interest in digital image processing. Recently, convolutional neural networks have achieved great success in terms of both reconstruction accuracy and computational performance for single-image super-resolution. However, one of the key properties of neural networks is their strong dependence on the training set. Due to this fact, we propose to use firstly the network dividing input images into several classes and then use the different super-resolution network for each class.

Contact: a.morgacheva@g.nsu.ru

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

PERSISTENT AERIAL TRACKING FROM UAVS

Mueller M., Smith N., Ghanem B.

Abstract: We present a new aerial video dataset and benchmark for low altitude UAV target tracking and provide the first evaluation of many trackers on 123 new and fully annotated HD video sequences. Additionally, we have developed a photo-realistic UAV simulator that can be used to evaluate tracking algorithms in real-time scenarios, as well as, generate synthetic but photo-realistic tracking datasets with free ground truth annotations. Lastly, we propose a persistent, robust and autonomous object tracking system for unmanned aerial vehicles (UAVs). A novel strategy is employed to successfully track objects over a long period, by 'handing over the camera' from one UAV to another.

Contact: matthias.mueller.2@kaust.edu.sa

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

FROM BENEDICT CUMBERBATCH TO SHERLOCK HOLMES - CHARACTER IDENTIFICATION IN TV SERIES

Nagrani A., Zisserman A.

Abstract: We propose and implement a new approach to character recognition in TV series with supervision provided only by actor face images from the web. In doing so, we eschew the need for transcripts and subtitles. To adapt from the actor's face to the character's face, we employ an automated semi-supervised learning approach involving face context and speaker identification from raw audio. We correctly identify profile faces, extreme pose and occlusion; and surpass all previous results on the Casablanca dataset.

Contact: arsha@robots.ox.ac.uk

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

TOWARDS A VISUAL PRIVACY ADVISOR: UNDERSTANDING AND PREDICTING RISKS IN IMAGES

Orekondy T., Schiele B., Fritz M.

Abstract: We propose a Visual Privacy Advisor, an approach to enforce users' desired privacy settings on image content. We first create a dataset of 22k images, annotated with 68 privacy attributes. Second, we run a user study to understand privacy preferences w.r.t to these attributes. Third, we propose models to predict user-specific privacy scores from images. Our model outperforms judgment of users, who often fail to enforce their own privacy preferences.

Contact: orekondy@mpi-inf.mpg.de

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

A UNIFIED APPROACH OF MULTI-SCALE DEEP AND HAND-CRAFTED FEATURES FOR DEFOCUS ESTIMATION

Park Jinsun, Tai Yu-Wing, Cho Donghyeon, Kweon In So

Abstract: We introduce robust and synergetic hand-crafted features and a simple but efficient deep feature from a convolutional neural network (CNN) architecture for defocus estimation. This paper systematically analyzes the effectiveness of different features, and shows how each feature can compensate for the weaknesses of other features when they are concatenated. Experimental results show that our algorithm is superior to state-of-the-art algorithms in terms of defocus estimation.

Contact: zzangjinsun@gmail.com

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

VIRALIENCY: POOLING LOCAL VIRALITY

Pilzer A., Alameda-Pineda X., Xu D., Sebe N., Ricci E.

Abstract: Nowadays we are exposed to a huge amount of media content in our connected world. Virality is the quality of an image or video to be rapidly and widely spread in social networks. Therefore, automatic understanding of image virality will be a key feature for many applications like advertising. Recent advancements in computer vision, show the ability of fully convolutional neural networks to learn class related activation maps. Moreover, through global pooling it is straightforward to figure them. We extend this concept proposing a pooling layer that learns the size of the support area to average: the learned top-N average pooling (LENA). Then, we design an end-to-end trainable siamese fully-convolutional network for virality localisation and virality classification. We report experiments on two publicly available datasets annotated for virality and plot local virality, viraliency, maps.

Contact: andrea.pilzer@unitn.it

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

MULTIVIEW MACHINE LEARNING USING AN ATLAS OF CARDIAC CYCLE MOTION

Puyol Esther, Langet Helene, Schnabel Julia, De Craene Mathieu, Piro Paolo,
King Andrew P.

Abstract: Statistical cardiac motion atlases provide a valuable tool for the quantification and evaluation of heart disease. In this work, we propose a framework for building a multimodal cardiac motion atlas from MR and ultrasound data and incorporate a multiview classifier to exploit the complementary information provided by the two modalities. We demonstrate that our novel framework is able to detect non ischemic dilated cardiomyopathy patients from ultrasound data alone with 83.50% classification accuracy, whilst still exploiting the MR based information from the multimodal atlas.

Contact: esther.puyol_anton@kcl.ac.uk

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

USING LSTM TO REDUCE COMPUTATIONAL COST IN DEEP CONVOLUTIONAL NEURAL NETWORK

Ray A., Sakaue F., Sato J.

Abstract: In this work we propose a method for reducing the computational cost in Deep CNNs. In our method, an LSTM is used for preserving the essential layer-wise information extracted by the CNNs and acting as a bias on the final fully connected layer to correctly classify images with relatively less network size. We believe that this will help in decreasing the computational cost and make real time applications with Deep CNN possible in computationally constrained environments.

Contact: anirbanray@cv.nitech.ac.jp

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

INTEGRATING STEREO VISION WITH A CNN TRACKER FOR A PERSON-FOLLOWING ROBOT

Sahdev R., Chen B.X., Tsotsos J.K.

Abstract: A Stereo vision based CNN Tracker for a person following robot is introduced. The robot follows the human in real time in a dynamic environment using a CNN which is trained online from scratch. Our Robot can follow the human under challenging situations - pose changes, illumination changes, appearance changes, wearing/removing a jacket/backpack, exchanging jackets, etc. The robot can follow the target even when the target is transiently not seen in the scene. A stereo dataset is also built.

Contact: raghavendersahdev@gmail.com

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

ENHANCING GRASSMANN DISCRIMINANT ANALYSIS

Souza L., Gatto B., Hino H., Fukui K.

Abstract: An enhanced algorithm of discriminant analysis on Grassmann manifold (eGDA) for image set classification is proposed. A limitation of GDA [1] is that its discriminative space is determined depending on the data points on the manifold, impairing its classification ability if the data are not so discriminative. To overcome this, we project subspaces onto a generalized difference subspace (GDS) [2] before manifold mapping. eGDA has been used for face, hand shape, and motion recognition [3, 4].

Contact: lincons@cvlab.cs.tsukuba.ac.jp

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

ONLINE SELF-SUPERVISED FCN FOR FREE-SPACE DETECTION WITH INTELLIGENT VEHICLES

Sanberg W., Dubbelman G., de With, P.

Abstract: A key component for Advanced Driver Assistance Systems (ADAS) is fast free-space detection to determine where the vehicle can drive. A fundamental challenge is that traffic scenes come in a wide variety (urban/rural, highway/city-center) under varying imaging conditions (good/bad weather, day/night). Therefore, we propose a system that is fast, flexible and robust. We achieve this by tuning a small FCN of around 1000 nodes online (while driving) on weak labels generated from disparity analysis.

Contact: w.p.sanberg@tue.nl

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

SURFACE INSPECTION OF OPTICAL ELEMENTS

Schöch A., Perez P., Linz-Dittrich S., Bach C., Ziolk C.

Abstract: In industry, manual visual inspection is typically applied to assess surface imperfections according to the standard DIN ISO 14997. We propose a machine vision setup to mimic the human testers inspection process. It consists of multiple cameras and LED light sources. By enabling individual LED sources, any movement during acquisition can be omitted. Results indicate a good agreement with the de facto standard manual inspection method from industry.

Contact: schoech@ntb.ch

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

DEEPWARP: PHOTOREALISTIC IMAGE RESYNTHESIS FOR GAZE MANIPULATION

Ganin Y., Kononenko D., Sungatullina D., Lempitsky V.

Abstract: In this work, we consider the task of generating highly-realistic images of a given face with a redirected gaze. We treat this problem as a specific instance of conditional image generation, and suggest a new deep architecture that can handle this task very well as revealed by numerical comparison with prior art and a user study. Our deep architecture performs coarse-to-fine warping with an additional intensity correction of individual pixels. All these operations are performed in a feed-forward manner, and the parameters associated with different operations are learned jointly in the end-to-end fashion. After learning, the resulting neural network can synthesize images with manipulated gaze, while the redirection angle can be selected arbitrarily from a certain range and provided as an input to the network.

Contact: diana.sungatullina@gmail.com

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

MATCHMOVING ON SET WITH THE USE OF REAL-TIME VISUAL-INERTIAL LOCALIZATION AND DEPTH ESTIMATION

Tarko J., Campbell N.

Abstract: We propose a new method of real-time camera tracking on set, using an innovative tablet that can localise itself by means of sensors bundle. Collected sensor data can be then used to create a virtual camera and to insert computer generated elements into a real video.

Contact: jkt26@bath.ac.uk

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

TWO-STREAM CONVOLUTIONAL NETWORKS FOR DYNAMIC TEXTURE SYNTHESIS

Tesfaldet Matthew., Brubaker Marcus A., Derpanis Konstantinos G.

Abstract: We introduce a two-stream model for dynamic texture synthesis based on pre-trained convolutional networks (ConvNets) that target two independent tasks: object recognition and optical flow prediction. Given an input dynamic texture, the object recognition ConvNet models the per-frame appearance of the input texture, while the optical flow ConvNet models its dynamics. To generate a novel texture, a noise sequence is optimized to match the feature statistics from each stream of the input texture.

Contact: mtesfald@eecs.yorku.ca

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

BEYOND BACTERIA ANALYSIS WITH DEEP LEARNING AND HYPERSPECTRAL DATA

Turra G., Arrigoni S., Signoroni A.

Abstract: Full Laboratory Automations are modifying the way microbiologists perform diagnostic tasks. In this work, we want to assess the feasibility of Deep Learning combined with hyperspectral analysis to identify bacteria colonies. From hyperspectral acquisitions in the VNIR range and spatial-spectral processing of urinary tract infections, we classify spectral signature of colonies with a Convolutional Neural Network, assessing its accuracy compared to conventional machine learning approaches.

Contact: g.turra004@unibs.it

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

ONLINE ADAPTATION OF CONVOLUTIONAL NEURAL NETWORKS FOR VIDEO OBJECT SEGMENTATION

Voigtlaender P., Leibe B.

Abstract: We tackle semi-supervised video object segmentation, i.e. segmenting the pixels of an object in a video using the ground truth pixel mask for the first frame. We build on the one-shot video object segmentation (OSVOS) approach which uses a pretrained network and fine-tunes it on the first frame. At test time OSVOS is not able to adapt to large changes in object appearance. To overcome this limitation, we propose Online Adaptive Video Object Segmentation (OnAVOS) which updates the network online.

Contact: voigtlaender@vision.rwth-aachen.de

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

ROBUST REGISTRATION FRAMEWORK OF MULTI-MODALITY MR DATA

Wang C, MacGillivray T, Koutraki Y, Boardman J, Sparrow S, Pataky R, and Semple S

Abstract: Combining the information obtained from multi-contrast MR data with other modalities by image registration is a useful first step for post-processing. Mutual information (MI) is a common technique used in many registration methods, calculated from the joint intensity histogram. However, common artefacts in MR images, for example, the intensity inhomogeneity shown in figure 1 may cause wrong calculation of histogram which may destroy registration performance. We proposed a framework based on ranged histogram specification and a novel particle swarm optimizer for robust 3D alignment of MR data. The framework is demonstrated using multi-contrast cardiovascular data acquired in Edinburgh Imaging Facility QMRI, University of Edinburgh.

Contact: chengjia.wang@ed.ac.uk

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

SILNET: 3D SHAPE FROM SILHOUETTES

Wiles O., Zisserman A.

Abstract: The objective of this work is 3D shape understanding from single and multiple images. SilNet learns this by generating the silhouette of an object at a new angle given a set of views.

SilNet handles an arbitrary number of images as input in an order-agnostic manner, and its performance improves with the addition of each new view at test time.

We test SilNet on ShapeNet and two new datasets: a blobby object dataset useful for pre-training and a sculpture dataset.

Contact: ow@robots.ox.ac.uk

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2

WHAT ARE YOU LOOKING AT IN A MUSEUM

Yang J., Sörös G.

Abstract: Museums and galleries can enhance the visitors' experience and deepen the understanding about the visitors' feelings by involving new technologies, such as smart audio guides, interactive guide robots, and eye-tracking systems. We apply eye trackers to record visitors' viewing behaviors in museum scenarios and leverage the results in three application domains: interests analysis, display form improvement, and smart guidance.

Contact: jing.yang@inf.ethz.ch

Presentation Type: Poster

Date: Tuesday 11 July 2017

Time: 21:30

Poster Session: 2