Demo Abstracts

Tuesday December 3, Morning

Demo 1:

Name: Li CHENG, Bioinformatics Institute, A*STAR, Singapore, chengli@bii.a-star.edu.sg

Associated paper: Our Paper with title "Efficient Hand Pose Estimation from a Single Depth Image" and id (1078) has been accepted into the main conference as a poster paper. In the paper, we tackle the practical problem of hand pose estimation from a single noisy depth image. Our approach is able to work with Kinect type noisy depth images, and reliably produces pose estimations of general motions efficiently. There are some further improvement we have made after the paper submission (which leads to improved results and faster rate). We are keen to demo to the ICCV audience our up-to-date system.

Full list of authors: Chi XU, and Li CHENG, both from Bioinformatics Institute, A*STAR, Singapore

Abstract: We tackle the practical problem of hand pose estimation from a single noisy depth image. A dedicated three-step pipeline is proposed: Initial estimation step provides an initial estimation of the hand in-plane orientation and 3D location; Candidate generation step produces a set of 3D pose candidate from the Hough voting space with the help of the rotational invariant depth features; Verification step delivers the final 3D hand pose as the solution to an optimization problem. We analyze the depth noises, and suggest tips to minimize their negative impacts on the overall performance. Our approach is able to work with Kinect-type noisy depth images, and reliably produces pose estimations of general motions efficiently (12 frames per second). Extensive experiments are conducted to qualitatively and quantitatively evaluate the performance with respect to the state-of-the-art methods that have access to additional RGB images. Our approach is shown to deliver on par or even better results.

Demo 2:

Name: Jürgen Sturm, Technical University of Munich, sturmju@in.tum.de

Associated paper: This work has been presented recently at the German Conference on Pattern Recognition (GCPR, formerly DAGM): CopyMe3D: Scanning and Printing Persons in 3D (J. Sturm, E. Bylow, F. Kahl, D. Cremers), In German Conference on Pattern Recognition (GCPR), 2013.

Paper:
http://vision.in.tum.de/_media/spezial/bib/sturm_etal_2013gcpr.pdf

Video: http://youtu.be/9B2IdD2jHGw

Full list of authors: Jürgen Sturm, Technical University of Munich; Erik Bylow, Lund University; Fredrik Kahl, Lund University; Daniel Cremers, Technical University of Munich

Abstract: http://youtu.be/9B2IdD2jHGw. We developed an approach to scan persons in 3D using a Kinect while they are rotating on a swivel chair. We post-process the scanned models automatically to close holes and to make it hollow, so that it can be printed cost-efficiently on a color 3D printer. Scanning a person takes around 10 seconds, the reconstruction runs in real-time so that the resulting 3D model is immediately visible on the computer. We can save these models to a VRML file so that the ICCV participants can take their 3D models home. We recently scanned persons live during a talk at the GCPR conference. We believe that this is a nice demo, as it showcases novel applications and challenges that arise from advances in 3D sensor and printer technology. We will also bring some of the 3D models along to demonstrate how the final result looks like.

Demo 3:

Name: Jakob Engel, Technical University Munich, jajuengel@gmail.com

Associated paper: Paper#: 1567 Title: Real-Time Semi-Dense Monocular SLAM: A Statistical Approach. Accepted as Poster.

Full list of authors: Jakob Engel, Jürgen Sturm, Daniel Cremers all from Technical University Munich

Abstract: We propose a statistical approach to perform real-time, semi-dense monocular SLAM on a single CPU. The key idea is to continuously estimate a semi-dense inverse depth map associated with each camera frame and to in turn use this depth map to track the motion of the camera. More specifically, we estimate for all pixels which have a non-negligible image gradient a Gaussian-distributed hypothesis for the inverse depth at that location. We propagate this inverse depth information over time and update it with stereo observations from images which provide the optimal baseline for the given pixel, allowing to accurately estimate the depth for both, far-away and close-by regions. In terms of tracking accuracy and computational speed, the proposed method compares favorably to both state-of-the-art dense and feature-based SLAM algorithms. We believe that this real-time capable CPU-based mono-SLAM approach is of major practical value for numerous applications such as the autonomous navigation of flying robots.

Tuesday December 3, Afternoon

Demo 4:

Name: Lorenz Meier, ETH Zurich, lm@inf.ethz.ch

Associated paper: ICCV Paper Submission #1725: Live Metric 3D Reconstruction on Mobile Phones

Full list of authors: Petri Tanskanen, Kalin Kolev, Lorenz Meier, Federico Camposeco, Olivier Saurer and Marc Pollefeys, all authors are affiliated to: Computer Vision and Geometry Group, ETH Zurich, Switzerland

Abstract: In this paper, we propose a complete on-device 3D reconstruction pipeline for mobile monocular hand-held devices, which generates dense 3D models with absolute scale on-site while simultaneously supplying the user with real-time interactive feedback. The method fills a gap in current cloud-based mobile reconstruction services as it ensures at capture time that the acquired image set fulfills desired quality and completeness criteria. In contrast to existing systems, the developed framework offers multiple innovative solutions. In particular, we investigate the usability of the available on-device inertial sensors to make the tracking and mapping process more resilient to rapid motions and to estimate the metric scale of the captured scene. Moreover, we propose an efficient and accurate scheme for dense stereo matching which allows to reduce the processing time to interactive speed. We demonstrate the performance of the reconstruction pipeline on multiple challenging indoor and outdoor scenes of different size and depth variability.

Demo 5:

Name: Shicheng Zheng, The Chinese University of Hong Kong, zsc.leigh@gmail.com

Associated papers:

[1] Unnatural L0 Sparse Representation For Natural Image Deblurring, CVPR 2013

[2] Two-Phase Kernel Estimation for Robust Motion Deblurring, ECCV 2010

Full list of authors: Shicheng Zheng, The Chinese University of Hong Kong; Li Xu, The Chinese University of Hong Kong

Abstract: We build a fast and robust deblurring software incorporating latest powerful blind deconvolution technologies to remove blur. It is capable of handling large blur kernel and recovering subtle structures and fine details with fast CPU implementation. GPU is also enabled.

Demo 6:

Name: Guofeng Zhang, State Key Lab of CAD&CG, Zhejiang University, zhangguofeng@cad.zju.edu.cn

Associated papers:

[1] Wei Tan, Haomin Liu, Zilong Dong, Guofeng Zhang and Hujun Bao. Robust Monocular SLAM in Dynamic Environments. International Symposium on Mixed and Augmented Reality (ISMAR), 2013.

[2] Zilong Dong, Guofeng Zhang, Jiaya Jia, and Hujun Bao. Efficient Keyframe-Based Real-Time Camera Tracking. Computer Vision and Image Understanding, accepted.

[3] Guofeng Zhang, Zilong Dong, Jiaya Jia, Tien-Tsin Wong, and Hujun Bao. Efficient Non-Consecutive Feature Tracking for Structure-from-Motion. European Conference on Computer Vision (ECCV), 2010.

Full list of authors: Guofeng Zhang, State Key Lab of CAD&CG, Zhejiang University; Haomin Liu, State Key Lab of CAD&CG, Zhejiang University; Zilong Dong, State Key Lab of CAD&CG, Zhejiang University; Wei Tan, State Key Lab of CAD&CG, Zhejiang University; Jiaya Jia, The Chinese University of Hong Kong; Hujun Bao, State Key Lab of CAD&CG, Zhejiang University

Abstract: We propose a robust monocular SLAM system which can work in challenging environments. A novel keyframe-based global localization method is proposed, which requires an offline module to select features from a group of reference images and constructs an optimal set of keyframes to approximately cover the entire space. In order to alleviate the effort of offline preprocessing, an online reference map extension module is proposed to realtime reconstruct new 3D features and select online keyframes to extend the keyframe set. Our system also can handle the dynamic scenes. Our online keyframe representation and updating method can adaptively model the dynamic environments, where the appearance or structure changes can be effectively detected and handled. The keyframes with large changed areas will be replaced by newly selected frames. In addition, a novel prior-based adaptive RANSAC algorithm (PARSAC) is proposed to efficiently remove outliers even when the inlier ratio is rather low. Experimental results demonstrate that the proposed system can robustly work in a large-scale scene with dynamic objects, and outperforms the state-of-the-art SLAM systems (e.g. PTAM).

Demo 7:

Date: December 3 (Afternoon)

Name: Tae-Kyun Kim, Imperial College London, tk.kim@imperial.ac.uk

Associated paper: ICCV 2013 oral paper: "Real-time Articulated Hand Pose Estimation using Semi-supervised Transductive Regression Forests"

Full list of authors: Danhang Tang (Imperial College London, London, UK), Tsz-Ho Yu (University of Cambridge, Cambridge, UK), Tae-Kyun Kim (Imperial College London, London, UK)

Abstract: This paper presents the first semi-supervised transductive algorithm for real-time articulated hand pose estimation. Noisy data and occlusions are the major challenges of articulated hand pose estimation. In addition, the discrepancies among realistic and synthetic pose data undermine the performances of existing approaches that use synthetic data extensively in training. We therefore propose the Semi-supervised Transductive Regression (STR) forest which learns the relationship between a small, sparsely labelled realistic dataset and a large synthetic dataset. We also design a novel data-driven, pseudo-kinematic technique to refine noisy or occluded joints. Our contributions include: (i) capturing the benefits of both realistic and synthetic data via transductive learning; (ii) showing accuracies can be improved by considering unlabelled data; and (iii) introducing a pseudo-kinematic technique to refine articulations efficiently. Experimental results show not only the promising performance of our method with respect to noise and occlusions, but also its superiority over state-of-the-arts in accuracy, robustness and speed.

Wednesday December 4, Morning

Demo 8:

Name: David Ferstl, Institute for Computer Graphics and Vision, Graz University of Technology, ferstl@icg.tugraz.at

Associated paper: In this demo we try to show the practical application of our accepted ICCV paper: Image Guided Depth Upsampling using Anisotropic Total Generalized Variation.

Full list of authors: David Ferstl - Institute for Computer Graphics and Vision / Graz University of Technology; Christian Reinbacher - Institute for Computer Graphics and Vision / Graz University of Technology; Rene Ranftl - Institute for Computer Graphics and Vision / Graz University of Technology; Matthias Rüther - Institute for Computer Graphics and Vision / Graz University of Technology; Horst Bischof - Institute for Computer Graphics and Vision / Graz University of Technology

Abstract: In this work we present a novel method for the challenging problem of depth image upsampling. Modern depth cameras such as Kinect or Time of Flight cameras deliver dense, high quality depth measurements but are limited in their lateral resolution. To overcome this limitation we formulate a convex optimization problem using higher order regularization for depth image upsampling. In this optimization an anisotropic diffusion tensor, calculated from a high resolution intensity image, is used to guide the upsampling. We derive a numerical algorithm based on a primal-dual formulation that is efficiently parallelized and runs at multiple frames per second.

Demo 9:

Name: Ralf Dragon, ETH Zurich, dragon@vision.ee.ethz.ch

Associated papers:

[1] Ralf Dragon, Jörn Ostermann, Luc Van Gool, "Robust Realtime Motion-Split-And-Merge for Motion Segmentation", GCPR (DAGM), 2013

[2] Ralf Dragon, Bodo Rosenhahn and Jörn Ostermann, "Multi-Scale Clustering of Frame-to-Frame Correspondences for Motion Segmentation", ECCV 2012

Full list of authors: Ralf Dragon, Luc Van Gool (Computer Vision Lab, ETH Zurich)

Abstract: Real-Time Motion Segmentation Motion segmentation is the grouping of trajectories according to their motion over time. In the past years, it has been receiving increasing attention since it can be used as a strong prior in dense object segmentation, for the unsupervised learning of object detectors, or for tracking. However, although more than a decade has passed since the first methods were proposed, most approaches are still not suitable for real-world applications as they are either limited by slow processing, they need the number of motions as prior knowledge, or they are not able to handle incomplete or erroneous trajectories. In our ECCV 2012 paper, we presented multi-scale motion clustering as a two-step approach which overcomes these problems and which performs very good on standard benchmarks. It consists of many frame-to-frame motion segmentations which are combined in one final clustering step. After a concise analysis of the frame-to-frame step in our DAGM 2013 paper, we were able to reduce the runtime and the error a lot which enables robust and realtime motion segmentation. In our live demo, we carry out motion segmentation on consecutive blocks. The result is displayed immediately such that an interactive feedback is provided.

Demo 10:

Name: Akshay Asthana, Imperial College London, a.asthana@imperial.ac.uk

Associated paper: Robust Discriminative Response Map Fitting with Constrained Local Models. Akshay Asthana, Stefanos Zafeiriou, Shiyang Cheng and Maja Pantic. In Proc. of 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2013), Portland, Oregon, USA, June 2013.

http://www.doc.ic.ac.uk/~aasthana/CVPR2013/AAsthana_CVPR2013.pdf

Demo videos available at:

https://sites.google.com/site/akshayasthana/clm-wild-code

Full list of authors: Akshay Asthana (Imperial College London); Stefanos Zafeiriou (Imperial College London); Shiyang Cheng (Imperial College London); Jie Shen (Imperial College London); Maja Pantic (Imperial College London and University of Twente)

Abstract: We present a fully-automatic and real-time facial landmark point tracking system that can robustly handle uncontrolled imaging conditions. The system can detects 66 facial landmark points and estimate the 3D head-pose at over 30fps on Intel Xeon 3.80 GHz processor with NVIDIA GeForce GTX 660 (960 CUDA cores). The system is based on the novel discriminative fitting of response map framework, proposed by Asthana et al (CVPR 2013), which has shown state-of-the-art performance for the task of facial landmark detection on multiple uncontrolled databases. The novelty lies in the use of texture model based on the response maps generated via discriminantly trained filters (i.e. patch- experts). Another benefit of this framework is that it allows for the direct use of 3D shape model during the alignment procedure and therefore, can be easily used for applications that require 3D facial information (for example, pose normalization for face recognition or facial expression recognition, facial animation, HCI interfaces etc). Therefore, using the current facial landmark point tracking system, we also show that the applications such as facial expression recognition and facial performance transfer can be performed in real-time under uncontrolled imaging conditions. The software will be made public at

http://ibug.doc.ic.ac.uk/resources.

Wednesday December 4, Afternoon

Demo 11:

Name: Markos Sigalas, Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH), msigalas@ics.forth.gr

Associated papers:

[1] Robust Model-based 3D Torso Pose Estimation in RGB-D sequences. Sigalas M., Pateraki M., Oikonomidis I., Trahanias P. (submitted to 2nd IEEE Workshop on Dynamic Shape Capture and Analysis, to be held within ICCV 2013)

[2] Model-based 3D Torso Pose Estimation from RGB-D data. Sigalas M., Pateraki M., Trahanias P. Demo Session in Computer Vision Pattern Recognition (CVPR), June 25-27, 2013.

[3] Visual tracking of hands, faces and facial features of multiple persons. Baltzakis H., Pateraki M., Trahanias P., 2012. Machine Vision and Applications. [doi: http://dx.doi.org/10.1007/s00138-012-0409-5]

[4] Using Dempster's rule of combination to robustly estimate pointed targets. Pateraki M., Baltzakis H., Trahanias P., 2012. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 14-18 May, St. Paul, Minnesota, USA. [doi: http://dx.doi.org/10.1109/ICRA.2012.6224870]

[5] Visual estimation of pointed targets for robot guidance via fusion of face pose and hand orientation. Pateraki M., Baltzakis H., Trahanias P., 2011. In Proc. of the 1st IEEE Workshop on Challenges and Opportunities in Robot Perception, held within the 13th International Conference on Computer Vision (ICCV), 6-13 November, Barcelona, Spain. [doi: http://dx.doi.org/10.1109/ICCVW.2011.6130368]

[6] Gesture recognition based on arm tracking for human-robot interaction. Sigalas M., Baltzakis H., Trahanias P. In Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Taipei, Taiwan, October 2010. [doi: http://dx.doi.org/10.1109/IROS.2010.5648870]

Full list of authors: Markos Sigalas(1,2), Maria Pateraki(1), Panos Trahanias(1,2)

1 Institute of Computer Science (ICS) - Foundation for Research and Technology - Hellas (FORTH), Greece

2 Department of Computer Science, University of Crete, Greece

Abstract: The extraction of pose related information from the human body is a challenging computer vision task and of utmost importance in free-form Human Robot Interaction in naturalistic environments. In this context the proposed demo presents our implemented system for extracting upper body pose, utilizing information from commercial RGB-D Sensors. A model-based method for the estimation of the upper body configuration, including torso and arms, has been developed and implemented under the assumption that no initialization phase is possible in order to commence interaction, and that pose recovery and tracking should remain unaffected from partial occlusions. The segmentation of the body of each user relies on our previous work on face identification. Then the formulation of the "minimum projection criterion" is employed to estimate the specific user top-view that is aligned with the main axis of the human torso, thus giving rise to robust 3D torso orientation. Interestingly, the mentioned top-view re-projection of each segmented user can also effectively cope with partial occlusions. The extracted torso pose provides estimations of the 3D shoulder locations. The latter are used, along with detected hypotheses of hands, in order to extract arm configurations for each user via Particle Filtering.

Demo 12:

Name: Yin Li, Georgia Institute of Technology, yli440@gatech.edu

Associated papers:

Yin Li, Alireza Fathi, James M. Rehg. Learning to Predict Gaze in Egocentric Video, ICCV 2013

Alireza Fathi, Yin Li, James M. Rehg. Learning to Recognize Daily Actions using Gaze, ECCV 2012

Zhefan Ye, Yin Li, Alireza Fathi, Yi Han, Agata Rozga, Gregory D. Abowd, James M. Rehg. Detecting Eye Contact using Wearable Eye-Tracking Glasses, 2nd International Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction (PETMEI) in conjunction with UbiComp 2012

Full list of authors: Yin Li, Zhefan Ye, Alireza Fathi, James Mathew Rehg (School of Interactive Computing, College of Computing, Georgia Institute of Technology)

Abstract: With the recent advent of wearable cameras, there has been an increasing interest in egocentric vision. Egocentric vision aims at the automatic analysis of video captured from a first-person perspective. We plan to demonstrate a real-time system for detecting eye contact, an important aspect of face-to-face interactions. With a single pair of video-recording glasses, our method detects moments of eye contact between the camera wearer and a second person during naturalistic social interactions. Our system highlights a set of key functions in egocentric vision, including gaze tracking, ego-motion estimation, face detection/tracking and facial analysis. Our real-time implementation utilizes gaze tracking technology from SMI eye-tracking glasses to determine the camera wearer's point of gaze. We combine the egocentric video from the glasses and the face analysis technology by Omron OKAO face SDK to estimate the gaze direction of others. We identify moments of eye contact as the event of simultaneous, mutual looking at faces. Our demo suggests the proposed egocentric vision platform as a promising vehicle for understanding the camera wearer's behavior.

Demo 13:

Name: Federico Pernici, MICC University Of Florence, pernici@dsi.unifi.it

Associated paper:
http://www.micc.unifi.it/pernici/index_files/eccv2012demo.pdf

Full list of authors: Federico Pernici and Alberto Del Bimbo

Abstract: Back to Back Comparison of Long Term Tracking Systems

Recently several comparisons of tracking systems have been presented in [1] [2] and others are soon to be published in the workshop VOT2013 (in conjunction with the ICCV2013) and in the website: www.alov300.org which consists of more than 300 sequences. In particular the website reports several experiments showing that none of the current methods is able to track objects in the long term. This demonstrates an evident bias on how performance are so far computed. Indeed most of the works on object tracking focused on nuisance factors such: occlusion, lighting, pose, blur, excluding time. According to this, we are interested in showing some of the best long term tracking methods (as for example [3], ([4] demo authors), [5]) simultaneously compared in real time on a laptop (2 or 3 systems will run in parallel on the same machine). This will help participants understanding the pros and cons of the various approaches.

[1] S. Salti, A. Cavallaro, and L. Di Stefano. Adaptive appearance modeling for video tracking: Survey and evaluation. TIP2012.

[2] Y. Wu, J. Lim, , and M.-H. Yang. Online object tracking: A benchmark. CVPR2013.

[3] S. Hare, A. Saffari, and P. Torr. Struck: Structured output tracking with kernels. ICCV2011.

[4] FaceHugger: The ALIEN Tracker Applied to Faces. ECCV2012 Demo Session.

[5] Kalal, Zdenek, Krystian Mikolajczyk, and Jiri Matas. "Tracking-learning-detection." Pattern Analysis and Machine Intelligence, IEEE Transactions on 34.7 (2012): 1409-1422.

Thursday December 5, Morning

Demo 14:

Name: William Nguatem, Bundeswehr University Munich, william.nguatem@unibw.de

Associated paper: Nguatem, W., Drauschke, M. and Mayer, H., ROOF RECONSTRUCTION FROM POINT CLOUDS USING IMPORTANCE SAMPLING., Workshop (CMRT13) - City Models, Roads and Traffic 2013, Antalya

Full list of authors: Nguatem, William., Drauschke, Martin. and Mayer, Helmut (Bundeswehr University Munich, Institute of Applied Computer Science)

Abstract: We propose a novel fully automatic technique for roof fitting in 3D point clouds based on sequential importance sampling (SIS). Our approach makes no assumption of the nature (sparse, dense) or origin (LIDAR, image matching) of the point clouds and further distinguishes, automatically, between different basic roof types based on model selection. The algorithm comprises an inherent data parallelism, the lack of which has been a major drawback of most Monte Carlo schemes. A further speedup is achieved by applying a coarse to fine search within all probable roof configurations in the sample space of roofs. The robustness and effectiveness of our roof reconstruction algorithm is illustrated for point clouds of varying nature.

Demo 15:

Name: Andrew Comport, CNRS-I3S, University of Nice Sophia-Antipolis, Andrew.Comport@cnrs.fr

Associated papers:

[1] M. Meilland, T. Drummond. and A. I. Comport. "A Unified Rolling Shutter and Motion Model for Dense 3D Visual Tracking". International Conference on Computer Vision. 2013.

[2] M. Meilland and A. I. Comport. "Super-resolution 3D Tracking and Mapping". IEEE International Conference on Robotics and Automation. 2013.

[3] M. Meilland, C. Barat and A. I. Comport. "3D High Dynamic Range Dense Visual SLAM and Its Application to Real-time Object Re-lighting". International Symposium on Mixed and Augmented Reality. 2013.

Full list of authors: M. Meilland, I3S/CNRS, University of Nice Sophia Antipolis France; T. Drummond, Electrical and Computer Systems Engineering, Monash University, Australia; A.I. Comport, I3S/CNRS, University of Nice Sophia Antipolis France.

Abstract: We will demonstrate our live real-time dense localisation and mapping approach that unifies two different representations commonly used to define dense models. On one hand, much research has looked at 3D dense model representations using voxel grids in 3D. On the other hand, image-based key-frame representations for dense environment mapping have been developed. The proposed demonstration will show a new model which unifies various concepts and exhibits the main advantages of each approach within a common multi-keyframe framework. One of the main results of the proposed approach is its ability to perform large scale reconstruction accurately at the scale of mapping a building in real-time. The demonstrator also includes our work on super-resolution localisation and mapping, high dynamic range 3D mapping at large scales along with fast visual SLAM while modelling rolling shutter and motion blur deformations.

Demo 16:

Name: Frank Steinbrücker, Technical University of Munich, frank.steinbruecker@in.tum.de

Associated paper: Paper No. 1688: "Large-Scale Multi-Resolution Surface Reconstruction from RGB-D Sequences".

Full list of authors: Frank Steinbrücker, Jürgen Sturm, Daniel Cremers (Technical University of Munich)

Abstract: We demonstrate a system for online reconstruction of large environments. We track the camera by aligning the dense images of a standard RGB-D sensor and perform online loop-closures by pose graph optimization. For the dense 3D reconstruction of the geometry and texture, we combine the RGB-D images in a volumetric signed distance function. The system runs in real-time on a standard laptop computer and it is capable of preserving fine details in large reconstruction volumes by embedding the signed distance function in a multi-scale octree data structure.

Thursday December 5, Afternoon

Demo 17:

Name: Johannes Jordan, University of Erlangen-Nuremberg, johannes.jordan@cs.fau.de

Associated papers:

[1] J. Jordan, E. Angelopoulou. “Mean-shift Clustering for Interactive Multispectral Image Analysis”, IEEE International Conference on Image Processing, September 2013

[2] J. Jordan, E. Angelopoulou. “Hyperspectral Image Visualization With a 3-D Self-organizing Map”, IEEE 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, June 2013

[3] J. Jordan, E. Angelopoulou. “Supervised Multispectral Image Segmentation with Power Watersheds”, IEEE International Conference on Image Processing, September 2012, pp. 1585-1588.

[4] J. Jordan, E. Angelopoulou. “Gerbil - A Novel Software Framework for Visualization and Analysis in the Multispectral Domain”, VMV 2010: Vision, Modeling and Visualization, November 2010, pp. 259-266.

Full list of authors: Johannes Jordan, University of Erlangen-Nuremberg; Elli Angelopoulou, University of Erlangen-Nuremberg

Abstract: Multispectral imaging is an important tool for better understanding of image formation and reflectance phenomena. Wideband RGB data is not sufficient to draw meaningful interpretations from the captured data; instead, a significant amount of filter bands needs to be available. Research on computer vision methods that interpret or rely on the reflectance in the scene often relies on analyzing multispectral images. However, due to the high dimensionality of the data, both human observers as well as computers have difficulty interpreting this wealth of information.

In this demonstration, we show how analysis of a multispectral image can be conducted interactively by the researcher with a powerful visual framework. Our software 'Gerbil' incorporates a new paradigm for visual assistance of multispectral analysis that specifically addresses the lack of a seamless integration of spectral distribution and topology. It puts emphasis on the spectral gradient, which is shown to provide enhanced information for many reflectance analysis tasks. It also includes a rich toolbox for evaluation of image segmentation, global clustering and false-color display algorithms in the multispectral domain.

We demonstrate how several specific research interests in scene reflectance can be tackled in a simple workflow on captured image data. The software is open-source (http://gerbilvis.org/) and researchers are encouraged to apply it on their own analysis tasks.

Demo 18:

Name: Nabil Belbachir, AIT Austrian Institute of Technology, nabil.belbachir@ait.ac.at

Associated papers:

[1] A.N. Belbachir, R. Pflugfelder and R. Gmeiner, “A Neuromorphic Smart Camera for Real-time 360° Distortion-free Panoramas,” IEEE Conference on Distributed Smart Cameras, USA, 2010.

[2] A.N. Belbachir, M. Mayerhofer, D. Matolin and J. Colineau, “Real-time Panoramic Views Using BiCa360, the fast Rotating Dynamic Vision Sensor to up to 10 Rotations per Sec,” IEEE ISCAS2012 International Symposium on Circuits and Systems, South Korea, 2012.

Patents:

[3] A.N. Belbachir and R. Pflugfelder, “Mechanism and approach for the acquisition of panoramic views (“360° Panorama visual scanner“), Austrian patent AT507.543. (15.10.2010); European patent application EP 2182720A2

[4] A.N. Belbachir and M. Mayerhofer, “Mechanism and approach for Real-time 360° 3D Panoramic Views (“Rotating Stereo Dynamic Line Vision Sensor“), German patent application DE 102011120718.3 (12. 12.2011)

Full list of authors: Manfred Mayerhofer, AIT Austrian Institute of Technology

Abstract: This live demonstration aims to show a new smart HDR camera TUCO-3D providing real-time 360° panoramic stereo views by scanning the surroundings horizontally at a speed up to 10 rotations per seconds. The demonstration will show event data generated from the rotating dynamic vision sensor (Contrast sensitive) and rendered in 360° panoramic views (10 panoramas per sec). Furthermore, we will show the corresponding intensity images, which are reconstructed by an on board software in real-time for up 10 panoramas /sec. In a further demonstration anaglyph representations of the panorama will be displayed, which show the 3D views by using specific glasses (anaglyph), which will be made available to the visitors. Panoramic 3D vision is can be very fascinating to the visitors of our demo desk as it is not frequently used in consumer technologies. However this technology is gaining important for several applications like wide area surveillance and robotics. The visitors will have the opportunity to learn about the technology behind the camera and interact for new experiments.

Demo 19:

Name: Lixin Fan, Nokia Research Center, fanlixin@ieee.org

Associated paper: No associated papers.

Full list of authors: Lixin Fan, Junsheng Fu, Yu You, Kimmo Roimela, Petri Piippo and Ville-Veikko Mattila

Abstract: Mobile phone industry has witnessed tremendous growth of embedded sensory units including GPS, accelerometer, gyros and, most importantly, multi-megapixel camera. Huge amount of photographs & video clips are captured by nowadays mobile phone cameras, and consequently, it remains a challenging task to present these media data in an intuitive and user-friendly manner.

By exploiting 3D reconstructed scene as well as camera pose information, “Photo tourism” system provided an intriguing user experience for browsing large collections of photographs. One of the limitations of the system, however, is its incapability to register photos that belong to parts of the scenes disconnected from the reconstructed ones. This demo presents our research endeavor to improve 3D photo browsing user experiences by combining user created photographs with large scale Nokia HERE 3D map data. Using the map street-view images, our system is able to register disconnected photos that cannot be handled by PhotoToursim system. Furthermore, since reconstructed scenes and camera poses are now placed in a global map context, rendering photographs together with 3D map data e.g. Terrain and textured 3D building models leads to a novel photo navigation experience on mobile devices.

Demo 20:

Name: Lixin Fan, Nokia Research Center, fanlixin@ieee.org

Associated paper: No associated papers.

Full list of authors: Yu You, Lixin Fan, Alain Boyer, Tuomas Kantonen and Ville-Veikko Mattila

Abstract: Augmented reality applications are gaining popularity due to increased capabilities of modern mobile devices. Creating AR content however is tedious and traditionally done on desktop environments by professionals, with extensive knowledge and/or even programming skills required. In this demo, we demonstrate a complete mobile approach for creating vision-based AR in both indoor and outdoor environment. Using hyperlinks, Web mashups are built to dynamically augment the physical world by normal users without programing skills.

This booklet was designed by Jian Zhang, Miao Miao Liu, Jingsong Xu and Richard Hartley.