Overview of the Technical Program

Invited talk: "Rome wasn’t built in one day, but one day Virtual Rome will be" by Luc van Gool
June, 18 (Wednesday) - 09:00 to 10:00

Recognition
June, 18 (Wednesday) - 10:30 to 12:00
Fast Simultaneous Tracking and Recognition Using Incremental Keypoint Matching
Jonathan Mooser, Quan Wang, Supe Yu and Ulrich Neumann (USC, USA)

Visual Word based Location Recognition in 3D Models using Distance Augmented Weighting
Friedrich Fraundorfer (ETH-Zurich, Switzerland), Changchung Wu (UNC-Chapel Hill, USA), Jan-Michael Frahm (UNC-Chapel Hill, USA) and Marc Pollefeys (ETH-Zurich, Switzerland and UNC-Chapel Hill, USA)

Expression-invariant Non-rigid 3D Face Recognition: A Robust Approach to Expression-aware Morphing
Farhad Al-Osaimi, Mohammed Bennamoun and Ajmal Mian (UWA, Australia)

Understanding
June, 18 (Wednesday) - 13:30 to 15:00
The Efficient Extension of Globally Consistent Scan Matching to 6 DoF
Don Borrman, Jan Elsberg, Kai Ungerling, Andreas Nüchter and Joachim Hertzberg (University of Osnabrück, Germany)

Think Globally, Cluster Locally: A Unified Framework for Range Segmentation
Gene Yu, Michael Grossberg, George Wolberg and Ioannis Stamos (CUNY, USA)

Estimation of Euler Characteristic from Volumetric Data
Konstantin Ryfkinon (UMass-Lowell, USA), Karen Daniels (UMass-Lowell, USA), Daniel Klein (UMass-Lowell, USA), Bradford Jones and Vincent Durante (UMass-Lowell, USA)

Learning
June, 18 (Wednesday) - 15:30 to 17:00
Optimized Camera Ranking Algorithms for Real-time Image Based Rendering of Endoscopic Image Data
Marco Wintert, Bernhard Then and Günther Greiner (University of Erlangen-Nuremberg, Germany)

Directional Associative Markov Network for 3-D Point Cloud Classification
Daniel Munoz, Nicole Vandapel and Martial Hebert (CMU, USA)

Automatic 3D Video Summarization: Key Frame Extraction from Self-Similarity
Peng Huang, Adrian Hilton and Jonathan Storck (University of Surrey, UK)

Posters and Demo
June, 18 (Wednesday) - 19:00 to 22:00
Temporal Styles for Time-Varying Volume Data
Jean-Paul Babacaruni (UB, Norway), Ivan Vida (UB, Norway), Torsten Müller (SFU, Canada) and Eduard Gröller (TU-Vienna, Austria)

A Virtual Reconstruction of the Entrance of the Ripoll Monastery
Isaac Besora (UPC, Spain), Pere Brunel (UPC, Spain), Marco Callieri (ISTI-CNR, Italy), Antoni Chica (UPC, Spain), Massimiliano Consoli (ISTI-CNR, Italy), Matteo Delépine (ISTI-CNR, Italy), Daniel Morales (UPC, Spain), Jordi Moya (UPC, Spain), Guido Ranzuglia (ISTI-CNR, Italy) and Roberto Scopigno (ISTI-CNR, Italy)

Fast Surface Reconstruction and Segmentation with Ground-Based and Airborne LIDAR Range Data
Matthew Carlberg, James Andrews, Peterer Gay and Avinui Zahier (UC-Berkeley, USA)

A Curvature-Driven Probabilistic Strategy for Transmission of Arbitrary 3D Meshes over Unreliable Networks
Irene Cheng (U of A, Canada), Lihuang Ying (U of A, Canada) and Konosuke Danishida (UPenn, USA)

Stride Scheduling for Time-Critical Collision Detection
Daniel Comming (DRI, USA) and Oliver Swardt (University of Roestock, Germany)

Capturing a Surface Light Field under Virtual Illumination
Greg Coombe, Jan-Michael Frahm and Anselmo Lastra (UNC-Chapel Hill, USA)

A Fully Automatic Approach for Human Recognition from Profile Images Using 2D and 3D Ear Data
Syed Islam, Mohammed Bennamoun, Ajmal Mian and Rossen Daviev (UWA, Australia)

Generalized Detection and Merging of Loop Closures for Video Sequences
Manfred Klopschitz (TU-Graz, Austria), Christoph Zach (UNC-Chapel Hill, USA), Arnold Inschauer (TU-Graz, Austria) and Dieter Schmalstieg (TU-Graz, Austria)

Accurate Camera Calibration and Correction Using Rigidity and Radial Alignment Constraints
Yonghao Liu (Queens University, UK), Ali Olsb (SLD Ltd, UK), Anthony Jakas (SLD Ltd, UK) and Longzhuang Li (TAMU, USA)

Graph-based Stereo Matching by Incorporating Monocular Cues
Xiangyu Ma and Hungtin Zha (PKU, China)

Fast View Interpolation from Stereo: Simpler can be Better
Nicolas Martin and Sébastien Roy (UDEM, Canada)

Filling Holes in 3D Meshes using Image Restoration Algorithms
Santiago Salamancab Mito (UNEX, Spain), Maria del Pilar Merchán García (UNEX, Spain), Emiliano Pueyo-Hernandez (UNED, Spain), Antonio Adin Oliver (UCLM, Spain) and Carlos Camada Somolinos (UNED, Spain)

Three-Dimensional Facial Imaging using a Static Light Screen and a Dynamic Subject
Robert McKeon and Patrick Flynn (UND, USA)

Geometric Calibration of a Structured Light System Using Circular Control Points
Jean-Noélan Dubel, Felix Rochette and Patrick Hobert (U. Laval, Canada)

Towards Real-time Stereo using Non-uniform Image Sampling and Sparse Dynamic Programming
Michel Sarkan and Klaus Diepold (TU-München, Germany)

Photometric Stereo via Computer Screen Lighting for Real-time Surface Reconstruction
Grant Schindler (GaTech, USA)

GPU rendering for autostereoscopic displays
Francois de Sottere (Université Paris-Est, France), Vincent Nuezik (Kalo University, Japan) and Venceslas Bri (Université Paris-Est, France)

Neural Networks for Arm Movement Prediction in CVEs
Fred Stakem and Ghassan AlRegib (GaTech, USA)
3D-Model view characterization using equilibrium planes
Adrien Theetten, Tarik Filali Ansary and Jean-Philippe Vandeborre (LIFL, France)

Using Markov Random Fields and Algebraic Geometry to Extract 3D Symmetry Properties
Yunfeng Su and Andrew Withee (UNC-Charlotte, USA)

An efficient and memory-conserving implementation of multi-view stereo for wide-area reconstruction
Xenophon Zabulis (ICS-FORTH, Greece), Nicolaos Grammalidis (ITI-CERTH, Greece) and Georgios D. Frazer (ICS-FORTH, Greece)

Fast and High Quality Fusion of Depth Maps
Christopher Zach (UNC-Chapel Hill, USA)

Efficient 3D Shape Acquisition and Registration Using Hybrid Scanning Data
Hongwei Zheng, Dietmar Saupe, Andreas Boshler and Peter Opuchlik (University of Konstanz, Germany)

Invited talk: “Adding Color Data to 3D Scanned Models” by Roberto Scopigno
June, 19 (Thursday) - 09:00 to 10:00

Light and Texture
June, 19 (Thursday) - 10:30 to 12:00

A Three-tier Hierarchical Model for Capturing and Rendering of 3D Geometry and Appearance from 2D Images
Martin Jagersand, Neil Birkbeck and Dana Czobora (U of A, Canada)

Smooth and non-smooth wavelet basis for capturing and representing light
Cameron Upright, Dana Czobora and Martin Jagersand (U of A, Canada)

Technology for Systems
June, 19 (Thursday) - 13:30 to 15:00

An Energy Formulation for Establishing the Correspondence Used in Projector Calibration
Marc-Antoine Droste (CNRC, Canada), Guy Godin (CNRC, Canada) and Sebastien Roy (UDEM, Canada)

Proxy-Based Compression of 2-1/2 D Structure of Dynamic Events for Tele-immersive Systems
Poopla Verlani and P.J. Narayanan (IIIT-Hyderabad, India)

Two and More Cameras
June, 19 (Thursday) - 15:30 to 17:00

MRF Stereo with Statistical Parameter Estimation
Shahid Huq, Andreas Koschan, Beerna Abidi and Mongi Abidi (UTK, USA)

Stereo matching using interior point methods
Anind Bhattacharya and Camillo J. Taylor (UPenn, USA)

Marker-less motion capture of skinned models in a four camera set-up using optical flow and silhouettes
Luca Ballan and Guido Maria Corthals (UniPD, Italy)

Invited talk: “Salient Visualization” by Amitabh Varshney
June, 20 (Friday) - 09:00 to 10:00

Invited Session
June, 20 (Friday) - 10:30 to 12:30

Challenges in wide-area structure-from-motion
Marc Pollefeys (ETH-Zurich, Switzerland and UNC-Chapel Hill, USA)

Title TBA
Drew Steddy (Microsoft Live Labs, USA)

Title TBA
Jana Kosecka (GMU and Google, USA)

Fast, Automated, 3D, Airborne Modeling of Large Scale Urban Environments
Avideh Zakhor (UC Berkeley, USA)

Reconstruction
June, 20 (Friday) - 14:00 to 16:00

3D Object Reconstruction with Heterogeneous Sensor Data
Li Guan (UNC-Chapel Hill, USA and ETH-Zurich, Switzerland), Jean-Sebastien Franco (LaBRI, France) and Marc Pollefeys (ETH-Zurich, Switzerland and UNC-Chapel Hill, USA)

L-Tangent Norm: A Low Computational Cost Criterion for Choosing Regularization Weights and its Use for Range Surface Reconstruction
Florent Brunet (CAMPAR, Germany and LAISMEA / LIAK, France), Adrien Bartoli (LAISMEA, France), Romy Malgouyres (LIAK, France) and Nasar Navab (CAMPAR, Germany)

GPU Accelerating Speeded-Up Robust Features
Timothy Ternusaty, Lisleyn Franch and John Heideman (ArgonST Inc., USA)

Single-Step Planar Surface Extraction from Range Images
Shamusangalam Suganthan, Sonya Coleman and Bryan Scottney (University of Ulster, UK)
Invited Talks

Rome wasn’t built in one day, but one day Virtual Rome will be

Luc van Gool

Currently, there is a surging interest in large-scale city modeling. This work can be divided according to multiple criteria, including present situation versus state in the past, visualisation applications versus cartographic surveying, using mobile mapping versus images mined from public repositories, etc. Our recent work in several of these areas is described, e.g. large-scale measurements of georeferenced 3D positions with high precision (typically within 5 cm at city scale), large-scale cultural heritage site reconstructions like the next generation Rome Reborn model (modeling the complete, ancient city of Rome), and planet-wide, automatic 3D modelling and annotation of monuments.

At the ESAT lab of the Katholieke Universiteit Leuven he is the head of a multidisciplinary group of 30 vision researchers, about 10 of whom are post-doc. With his VISICS team he is involved in several European research, exchange, and training projects IST, Eureka, Esprit. Currently he is coordinator of 1 European project. At the Image Science Group (BIWI) of ETH (Zurich, Switzerland) he is the head of 20 researchers, specialised in the fields of medical imaging, robot vision, remote sensing and image indexing. The BIWI group is involved in several European and Swiss projects. He is a member of the Program Committee of several major international computer vision conferences (ICCV, ECCV, ICPR, ECMAST, MFI, etc.) and regularly serves as reviewer of leading scientific journals in the filed of vision. In 1998 he co-founded the spin-off company Eyetronics, which is specialised in 3D modeling. In its first year of existence they won already several awards with its 3D technology, including the European IT prize 1998.

Adding Color Data to 3D Scanned Models

Roberto Scopigno

The easy construction of detailed and accurate 3D models is becoming a reality by the increasing diffusion of 3D scanning technology. The reduction in cost of the scanning devices and the increasing availability of good processing tools (including emerging open source solutions) makes 3D scanning an enabling technology for the construction of shape models. The talk will present the capabilities of this technology focusing mainly on a less consolidated area: how color or surface reflection characteristics could be sampled and associated with reconstructed 3D shape models. The different approaches proposed will be reviewed, giving more emphasis to the more practical solutions for both acquiring color or surface reflection and mapping those data efficiently on surface meshes. Some examples of the results of current projects, mainly in the Cultural Heritage field, will be shown.

Roberto Scopigno is a Research Director at ISTI-CNR, an Institute of the Italian National Research Council (CNR). He leads the Visual Computing Lab of ISTI-CNR. He is currently engaged in research projects concerned with multiresolution data modeling and rendering, 3D scanning, surface reconstruction, scientific visualization and applications to Cultural Heritage. Most of these activities have been funded by EU grants (EU IST projects). He published more than one hundred twenty papers in international refereed journals/conferences and gave invited lectures or courses at several international conferences. He was Co-Chair of international conferences (Eurographics 1999, Rendering Symposium 2002, WSCG 2004, Geometry Processing Symp. 2004, Eurographics 2008). Since 2001 he is Co-Editor in Chief of the Computer Graphics Forum Journal. He has been elected member of the Eurographics Executive Committee on 2001, and appointed Vice Chair of the association in 2003.

Salient Visualization

Amitabh Varshney

Recent advances in data acquisition and scientific simulations result in visualization datasets that can easily overwhelm human comprehension. To deal with the resulting glut of visual clutter one turns to the human visual system for inspiration. The human visual system deals with the natural complexity of the world around us by focusing retinal hardware and attention on what is most important. Indeed, the principles of visual saliency have long been used in art, illustration, and photography to convey varying levels of importance of the constituent elements. In this talk I shall overview examples of how principles of visual saliency have been used over time leading up to the present in visual communication and scientific discovery.

Amitabh Varshney is a Professor of Computer Science at the University of Maryland. Varshney's research has addressed challenges in interactive 3D visualization for large datasets by reconciling realism with interactivity through multiresolution techniques and high-performance computing. He has served as the papers co-chair for IEEE Visualization 2000 and 2001, program co-chair for IEEE Visualization 2005, and conference co-chair for IEEE Visualization 2006 and 2007. He has served on program committees of several conferences. During 1999-2003 he served on the editorial board of IEEE Transactions on Visualization and Computer Graphics. Varshney received the NSF CAREER award in 1995 and the IEEE Visualization Technical Achievement Award in 2004.
Recognition

10:30 to 12:00 - June, 18 (Wednesday)

Session chair: Frank Dellaert

Fast Simultaneous Tracking and Recognition Using Incremental Keypoint Matching
Jonathan Mooser, Quan Wang, Suya You and Ulrich Neumann (USC, USA)

We present a unified approach to object recognition and object tracking, combining local feature matching with optical flow. Like many traditional recognition algorithms, we implement recognition by matching detected image patches against a database of known objects. In our case, however, matching is performed incrementally, meaning that only a few keypoints are tested at each frame until the complete object is identified. Recognition and tracking thus proceed in real-time, even with high dimensional features and an arbitrarily large database. Central to our work is the system by which keypoint matching and optical flow mutually aid one another. Keypoint matching recognizes an object and estimates its pose in order to initialize tracking. Optical flow tracking, in turn, maintains the object pose over subsequent frames, discarding newly matched keypoints that do not fit with the current pose estimation. We demonstrate that this powerful combination provides robust, real-time recognition and tracking of multiple objects in the presence of scale and orientation changes as well as partial occlusion.

The paper will be presented by Jonathan Mooser

Visual Word based Location Recognition in 3D Models using Distance Augmented Weighting
Friedrich Fraundorfer (ETH-Zurich, Switzerland), Changchang Wu (UNC-Chapel Hill, USA), Jan-Michael Frahm (UNC-Chapel Hill, USA) and Marc Pollefeys (ETH-Zurich, Switzerland and UNC-Chapel Hill, USA)

For visual word based location recognition in 3D models we propose a novel distance-weighted scoring scheme. Matching visual words are not treated as perfect matches anymore but are weighted with the distance of the original SIFT feature vectors before quantization. To maintain the scalability and efficiency of vocabulary tree based approaches PCA compressed SIFT feature vectors are used instead of the original SIFT features. A different Eigenspace is computed for each vocabulary tree cell to benefit from the variance reduction as result of the partitioned SIFT feature space. Experiments show a significant improvement in retrieval quality by incorporating the distance with small costs in computational time and memory.

The paper will be presented by Friedrich Fraundorfer

Expression-invariant Non-rigid 3D Face Recognition: A Robust Approach to Expression-aware Morphing
Faisal Al-Osaimi, Mohammed Bennamoun and Ajmal Mian (UWA, Australia)

An automatic approach to the recognition of 3D face under any facial expression is presented. It has the capacity to differentiate between expression deformations, those which are caused by expressions and interpersonal disparities that are essential for recognition. When two facial scans are matched the expression deformations are morphed out before similarity measure calculation while the interpersonal disparities are retained. A combination of PCA and ICP is used in our system. The PCA subspaces built in such a way it models the facial expressions while it does not model the human face. The approach was applied on the FRGC v2.0 dataset and superior recognition performance was achieved. The verification rates at 0.001 FAR were 98.35% and 97.73% for scans under neutral and non-neutral expressions, respectively.

The paper will be presented by Faisal Al-Osaimi
The Efficient Extension of Globally Consistent Scan Matching to 6 DoF

Dorit Borrmann, Jan Elseberg, Kai Lingemann, Andreas Nüchter and Joachim Hertzberg (University of Osnabrueck, Germany)

Over ten years ago, Lu and Milios presented a probabilistic scan matching algorithm for solving the simultaneous localization and mapping (SLAM) problem with 2D laser range scans, a standard in robotics. This talk presents an extension to this GraphSLAM method. Our iterative algorithm uses a sparse network to represent the relations between several overlapping 3D scans, computes in every step the 6 degrees of freedom (DoF) transformation in closed form and exploits efficient data association with cached k-d trees. Our approach is based on Gaussian error modeling and leads to globally consistent 3D maps, precise 6D pose and covariance estimates, as demonstrated by various experimental results.

The paper will be presented by

Dorit Borrmann

Think Globally, Cluster Locally: A Unified Framework for Range Segmentation

Gene Yu, Michael Grossberg, George Wolberg and Ioannis Stamos (CUNY, USA)

Modern range scanners can capture the geometry of large urban scenes on an unprecedented scale. While the volume of data is overwhelming, urban scenes can be approximated well by parametric surfaces such as planes. Piecewise planar representation can reduce the size of the data dramatically. Furthermore, it is ideal for rendering and other high-level applications. We present a segmentation algorithm that extracts a piecewise planar function from a large range image. Many existing algorithms for large datasets apply planar criteria locally to achieve efficient segmentations. Our novel framework combines local and global approximants to guarantee truly planar components in the output. To demonstrate the effectiveness of our approach, we present an evaluation method for piecewise planar segmentation results based on the minimum description length principle. We compare our method to region growing on simulated and actual data. Finally, we present results on large scale range images acquired at New York's Grand Central Terminal.

The paper will be presented by

Gene Yu

Estimation of Euler Characteristic from Volumetric Data

Konstantin Rybnikov (UMass-Lowell, USA), Karen Daniels (UMass-Lowell, USA), Daniel Klain (UMass-Lowell, USA), Bradford Jones and Vincent Durante (UMass-Lowell, USA)

Determining the shape of a complex 3D object from point data is important in medical tomography, molecular biology, X-ray crystallography and other areas. Using integral geometry we developed, implemented and tested a family of efficient estimators for the Euler characteristic. Our algorithms work with “volumetric” input, i.e., with point samplings of 3D bodies, not just their boundaries. They have \( O(N \log N) \) time complexity for \( N \) points, based on sorting, but have a significant advantage in terms of speed and memory when compared to methods that approximate unknown bodies with simplicial complexes or splines. Due to the nature of functions evaluated by our algorithms (polynomials of degree at most \( n \) for \( \mathbb{R}^n \)), we have exceptional robustness with respect to floating point errors. Our programs can be used for practical inquiries, such as estimating the number of connected components in the input data and finding the number of spherical voids in a connected 3D body.

The paper will be presented by

Konstantin Rybnikov
Learning

15:30 to 17:00 - June, 18 (Wednesday)

Session chair: Jana Kosecka

Optimized Camera Ranking Algorithms for Real-time Image Based Rendering of Endoscopic Image Data

Marco Winter, Bernhard Then and Günther Greiner (University of Erlangen-Nuremberg, Germany)

Unstructured Lumigraph Rendering is a well-established and flexible image-based rendering technique to create novel views from a set of acquired images. However, the standard implementation may not provide a pleasing visualization when images are captured under suboptimal conditions. In our scenario, lightfield information is gained using images from endoscopes that observe the operating field during minimal-invasive surgery. Suboptimal lighting conditions and limited mobility of the endoscope lead to visual artefacts in the final visualization of the scene. To resolve these problems, we propose a set of new algorithms for computing the ranking criterions that are used by the Unstructured Lumigraph for visualization. Using these algorithms, we get satisfying results for visualization of endoscopic image data captured under the aforementioned conditions. To achieve real-time visualization, the vital parts of our algorithms have been implemented using OpenGL and GLSL, thus making use of the features and speed of modern graphics hardware.

The paper will be presented by Marco Winter

Directional Associative Markov Network for 3-D Point Cloud Classification

Daniel Munoz, Nicolas Vandapel and Martial Hebert (CMU, USA)

In this paper we address the problem of automated three dimensional point cloud interpretation. This problem is important for various tasks from environment modeling to obstacle avoidance for autonomous robot navigation. In addition to locally extracted features, classifiers need to utilize contextual information in order to perform well. A popular approach to account for context is to utilize the Markov Random Field framework. One recent variant that has successfully been used for the problem considered is the Associative Markov Network (AMN). We extend the AMN model to learn directionality in the clique potentials, resulting in a new anisotropic model that can be efficiently learned using the subgradient method. We validate the proposed approach using data collected from different range sensors and show better performance against standard AMN and Support Vector Machine algorithms.

The paper will be presented by Daniel Munoz

Automatic 3D Video Summarization: Key Frame Extraction from Self-Similarity

Peng Huang, Adrian Hilton and Jonathan Starck (University of Surrey, UK)

In this paper we present an automatic key frame selection method to summarise 3D video sequences. Key-frame selection is based on optimisation for the set of frames which give the best representation of the sequence according to a rate-distortion trade-off. Distortion of the summarization from the original sequence is based on measurement of self-similarity using volume histograms. The method evaluates the globally optimal set of key-frames to represent the entire sequence without requiring pre-segmentation of the sequence into shots or temporal correspondence. Results demonstrate that for 3D video sequences of people wearing a variety of clothing the summarization automatically selects a set of key-frames which represent the dynamics. Comparative evaluation of rate-distortion characteristics with previous 3D video summarization demonstrates improved performance.

The paper will be presented by Peng Huang
**Temporal Styles for Time-Varying Volume Data**

Jean-Paul Balabanian (UiB, Norway), Ivan Viola (UiB, Norway), Torsten Möller (SFU, Canada) and Eduard Gröller (TU-Vienna, Austria)

This paper introduces interaction mechanisms for conveying temporal characteristics of time-varying volume data based on temporal styles. We demonstrate the flexibility of the new concept through different temporal style transfer function types and we define a set of temporal compositors as operators on them. The data is rendered by a multi-volume GPU raycaster that does not require any grid alignment over the individual time-steps of our data nor a rectilinear grid structure. The paper presents the applicability of the new concept on different data sets from partial to full voxel alignment with rectilinear and curvilinear grid layout.

The paper will be presented by Jean-Paul Balabanian

---

**A Virtual Reconstruction of the Entrance of the Ripoll Monastery**

Isaac Besora (UPC, Spain), Pere Brunet (UPC, Spain), Marco Callieri (ISTI-CNR, Italy), Antoni Chica (UPC, Spain), Massimiliano Corsini (ISTI-CNR, Italy), Matteo Dellepiane (ISTI-CNR, Italy), Daniel Morales (UPC, Spain), Jordi Moyés (UPC, Spain), Guido Ranzuglia (ISTI-CNR, Italy) and Roberto Scopigno (ISTI-CNR, Italy)

In this paper we present a project which aimed at virtually reconstructing the impressive (7x11 m.) portal of the Ripoll Monastery, Spain. The monument was acquired using triangulation laser scanning technology, producing a dataset of more than 2000 range maps for a total of more than 1 billion triangles. All the steps of the entire project are described, from the acquisition planning to the final setup for the dissemination to the public. In particular, we show how time-of-flight laser scanning data can be used to obtain a speed up in the alignment process, and how, after model creation and imperfections repairing, an interactive and immersive setup gives the public the possibility to navigate and visualize the high detail representation of the portal.

The paper will be presented by Roberto Scopigno

---

**Fast Surface Reconstruction and Segmentation with Ground-Based and Airborne LIDAR Range Data**

Matthew Carlberg, James Andrews, Peiran Gao and Avideh Zakhor (UC-Berkeley, USA)

Advances in range measurement devices in recent years have opened up new opportunities and challenges for fast 3D modeling of large scale environments. Applications of such technologies include virtual walk and fly through, urban planning, disaster management, object recognition, training, and simulations. In this paper, we present a general framework for surface reconstruction and segmentation using partially ordered 3D point clouds composed of registered ground-based and airborne range and color data. Our algorithms can be applied to a large class of LIDAR data acquisition systems, where ground-based data is obtained as a series of scan lines. We develop an efficient and scalable algorithm that reconstructs surfaces and segments ground-based range data simultaneously. We also propose a new algorithm for merging ground-based and airborne meshes which exploits the locality of the ground-based mesh. We demonstrate the effectiveness of our results on data sets obtained by two acquisition systems.

The paper will be presented by Matthew Carlberg
A Curvature-Driven Probabilistic Strategy for Transmission of Arbitrary 3D Meshes over Unreliable Networks

Irene Cheng (U of A, Canada), Lihang Ying (U of A, Canada) and Kostas Daniilidis (UPenn, USA)

Packet loss affects the receiving quality of 3D meshes transmitted over unreliable networks. While some applications are able to tolerate higher loss, others may need to restrict the loss below a specified level. In this work we describe a curvature-driven probabilistic strategy to control the adverse impact of packet loss. Critical mesh features, with high curvature, like sharp edges and corners are allocated more bandwidth to increase the rate of their successful transmission. When the probability of visual degradation exceeds an acceptable level, a group of curvature indices is added to the transmission pipeline. The size of the indices is governed by three parameters: The mesh resolution, the minimum required quality and the tolerance. We incorporate this new strategy with an earlier interleaved transmission approach. Experimental results show that the reconstructed meshes using the integrated strategy have higher visual quality.

The paper will be presented by Irene Cheng

Stride Scheduling for Time-Critical Collision Detection

Daniel Coming (DRI, USA) and Oliver Staadt (University of Rostock, Germany)

We present an event-based scheduling method for time-critical collision detection that meets real-time constraints by balancing and prioritizing computation spent on intersection tests without starvation. Our approach tests each potentially colliding pair of objects at a different frequency, with unbounded temporal resolution. We show that believability is preserved by adaptively prioritizing intersection tests to reduce errors in collision detection, using information about the objects and scene. Through the combination of kinetic sweep and prune with stride scheduling we continuously interleave rendering, broad phase collision pruning, narrow phase intersection testing, and collision response. This approach accrues no per-frame overhead and allows interruption at any point in collision detection, including the broad phase.

The paper will be presented by Daniel S. Coming

Capturing a Surface Light Field under Virtual Illumination

Greg Coombe, Jan-Michael Frahm and Anselmo Lastra (UNC-Chapel Hill, USA)

Surface light fields can be used to render the complex reflectance properties of a physical object. One limitation is that they can only represent the fixed lighting conditions of the environment where the model was captured. If a specific lighting condition is desired, then there are two options: either use a combination of physical lights as an approximation, or capture a full 6D surface reflectance field and only use the portion that corresponds to the desired lighting. In this paper we present a method for capturing a surface light field using the virtual illumination from an environment map. We use a simple setup consisting of a projector, a camera, a pan-tilt unit, and tracking fiducials to recreate the desired lighting environment. To decrease noise and improve the quality of the capture under low- and high-dynamic range environment maps, we use an extended version of the multiplexed illumination algorithm. We show results from objects captured under different lighting environments.

The paper will be presented by Greg Coombe
A Fully Automatic Approach for Human Recognition from Profile Images Using 2D and 3D Ear Data

Syed Islam, Mohammed Bennamoun, Ajmal Mian and Rowan Davies (UWA, Australia)

Using ear shape as a biometric trait is one of the most recent trends in the biometric research communities. In this work, a fully automatic and fast technique based on the AdaBoost algorithm is used to detect the human ear from 2D and the corresponding 3D profile image. A modified Iterative Closest Point (ICP) algorithm is then used for matching the extracted ear shapes. The ICP algorithm is applied hierarchically first on a lower and then on a higher resolution meshes of 3D ear data. We obtain a rank one ear recognition rate of 93% on the profile images of the University of Notre Dame biometrics database. The proposed recognition approach does not require any manual intervention or sharp extraction of ear contour from the detected ear region. Moreover, the system performance does not rely on the presence of a particular feature of the ear.

The paper will be presented by Syed Mohammed Shamsul Islam

Generalized Detection and Merging of Loop Closures for Video Sequences

Manfred Klopschitz (TU-Graz, Austria), Christopher Zach (UNC-Chapel Hill, USA), Arnold Irshara (TU-Graz, Austria) and Dieter Schmalstieg (TU-Graz, Austria)

In this work we present a method to detect overlaps in image sequences, and use this information to integrate overlapping sparse 3D structure from video sequences. The additional temporal information of these images is used to increase robustness over single image pair matching. A scanline optimization problem formulation is used to compute the best sequence alignment using wide-baseline image matching techniques. Compared to a direct dynamic programming approach, the scanline optimization approach increases the robustness of sequence alignment for general relative motions. The proposed alignment method is employed to integrate sparse 3D models reconstructed from separate video sequences. In addition loop closures are detected. Consequently, the 3D modeling process from sequential image data can be split into fast sequence processing and subsequent global integration steps.

The paper will be presented by Manfred Klopschitz

Accurate Camera Calibration and Correction Using Rigidity and Radial Alignment Constraints

Yonghuai Liu (Aberystwyth University, UK), Ala Al-Obaidi (SLD Ltd, UK), Anthony Jakas (SLD Ltd, UK) and Longzhuang Li (TAMU, USA)

In this paper, we develop a novel method for camera calibration and correction. The novel method first employs a rigidity constraint from the rigid rotation matrix and the radial alignment constraint from the pin-hole camera model to estimate both camera intrinsic and extrinsic parameters with a closed-form solution without considering the camera distortion. Then the well-known Levenberg-Marquardt (LM) algorithm is employed to optimize the parameters of interest in two steps: the first step optimizes the first order radial distortion coefficient and the z component of the camera position using a partial image formation model, and the second step optimizes all the parameters of interest using a complete image formation model: 4 intrinsic, 7 extrinsic and 4 distortion parameters. The LM algorithm is initialized either as the parameters estimated so far or as zero. The optimization is achieved through minimizing the sum of the squared differences between the distorted projected 3D world control points and their given corresponding distorted image points. The distorted points are finally corrected using again the LM algorithm initialized by the distorted image points themselves, minimizing the squared difference between the distorted corrected point and the given distorted image point. The experimental results based on both synthetic data and real images show that the proposed algorithm produces promising camera calibration and correction results.

The paper will be presented by Yonghuai Liu
Graph-based Stereo Matching by Incorporating Monocular Cues
Xiangyin Ma and Hongbin Zha (PKU, China)

Stereo vision is one of the most intensive and challenging problems in computer vision. It makes use of stereo cues to extract 3D information from 2D images. Besides stereo cues, there are some statistical learning based approaches which exploit monocular cues to predict the underlying 3D structure. As for human vision, the amazing ability for 3D interpretation is based on the combination of these two kinds of cues. Therefore, in this paper, we make an attempt to incorporate monocular cues into the stereo matching system. A two-level graph is utilized to fuse the low-resolution monocular cues and high-resolution stereo cues together. Then the optimal labeling results are calculated via graph-cuts. The experiment results show that we can obtain more accurate disparity map than is possible using either monocular or stereo cues alone.

The paper will be presented by
Xiangyin Ma

Fast View Interpolation from Stereo: Simpler can be Better
Nicolas Martin and Sébastien Roy (UDEM, Canada)

In this paper, we propose to rely only on images to generate novel views, and recall why modeling of the complete scene is often too expensive in the context of view interpolation. We investigate ways to achieve view interpolation by mean of forward and backward mapping. We present situations in which each one requires less computations and gives better results. Contrary to what we might expect, very simple stereo algorithms can produce very convincing interpolation despite providing really bad disparity maps. We propose to explain this with a probabilistic model of depth discontinuities. We test this model on synthetic data created to fit real image statistics and compare with images widely used in stereo. In practice, forward and backward mapping methods can rely on simple stereo algorithms running in real time, to produce very good results. A sequence of real images was acquired to allow accurate comparison of interpolated images, and standard metrics are used to assess the quality.

The paper will be presented by
Nicolas Martin

Filling Holes in 3D Meshes using Image Restoration Algorithms
Santiago Salamanca Miño (UNEX, Spain), María del Pilar Merchán García (UNEX, Spain), Emiliano Pérez Hernández (UNED, Spain), Antonio Adán Oliver (UCLM, Spain) and Carlos Cerrada Somolinos (UNED, Spain)

This work describes a method for filling holes in a 3D mesh based on 2D image restoration algorithms. Since these algorithms need an image as input, the first stage of the method concerns a 3D to 2D transformation for a range image creation. The image restoration algorithms are applied to it. Once the image has been repaired, the inverse transformation 2D to 3D is performed and the repaired 3D surface recovered. To test the method, artificial holes have been generated on a set of 3D surfaces. The goodness of the results has been measured from the comparison between the 3D original surfaces and the 3D repaired ones. An evaluation with commercial software has been carried out to show the validity of the method. The image restoration algorithms have been applied to 3D cultural heritage modeling with good results.

The paper will be presented by
Emiliano Pérez Hernández
Three-Dimensional Facial Imaging using a Static Light Screen and a Dynamic Subject

Robert McKeon and Patrick Flynn (UND, USA)

Many commercially available 3D sensors suitable for face image capture employ either passive (or texture-assisted) stereo imaging or structured illumination with a moving stripe. Both of these techniques require a stationary subject. We describe an initial design and evaluation of a fixed-stripe, moving object 3D scanner designed for human faces. Our method of acquisition requires the subject to walk through a light screen generated by two laser line projectors. Triangulation and tracking yield a 3D image of the subject’s face from multiple images. To demonstrate the accuracy of our initial design, a small-scale facial recognition experiment was executed. In an experiment involving 13 subjects with 4 images per subject, we achieve 92.3% rank one recognition using an Iterative Closest Point (ICP) based matching method, demonstrating the feasibility of the technique.

The paper will be presented by Robert McKeon

Geometric Calibration of a Structured Light System Using Circular Control Points

Jean-Nicolas Ouellet, Félix Rochette and Patrick Hébert (ULaval, Canada)

We present a new geometric calibration method for a structured light system combining a projector with a camera, using a planar target with circular control points. By solely exploiting the mapping between projected conics, the proposed method is strictly geometric and provides unbiased camera to projector correspondences during its application. Such a geometric method does not rely on radiometric calibration. Moreover, the method consistently ensures uniform coverage of the working volume and automatically avoids interference between both the projected and the printed patterns on the calibration target.

The paper will be presented by Jean-Nicolas Ouellet

Towards Real-time Stereo using Non-uniform Image Sampling and Sparse Dynamic Programming

Michel Sarkis and Klaus Diepold (TU-München, Germany)

Acquiring the 3D mesh of a scene from stereo images is a major task in computer vision. It usually involves several steps including stereo matching and meshing. Unfortunately, the time required to generate the 3D mesh is time demanding due to the large amount of pixels to be processed. In this work, we propose a framework to accelerate the overall process. The key issue is to first reduce the number of pixels by approximating an image with a content adaptive mesh. The nodes of the mesh are sparse and they represent the non-uniform samples of the image. To benefit from the reduced set of pixels, we formulate a dynamic programming based stereo matching algorithm which computes the depth only at the sparse samples.

The paper will be presented by Michel Sarkis
**Photometric Stereo via Computer Screen Lighting for Real-time Surface Reconstruction**

Grant Schindler (GaTech, USA)

We introduce a method which uses the light emitted by a computer screen to illuminate an object such as a human face from multiple directions, simultaneously capturing images with a webcam in order to perform photometric stereo. Dominant eigenvectors of the captured images provide surface normals, which are integrated into a 3D surface using Gauss-Seidel relaxation. The system runs at 10 frames per second on a consumer laptop computer.

The paper will be presented by

Grant Schindler

---

**GPU rendering for autostereoscopic displays**

François de Sorbier (Université Paris-Est, France), Vincent Nozick (Keio University, Japan) and Venceslas Biri (Université Paris-Est, France)

In recent years, stereoscopic technology has advanced from stereoscopic to autostereoscopic displays. These latter family involves to display several views of a scene. In the case of real-time computer graphics images, the standard approach consists in rendering every view independently. This paper presents an alternative method to generate multiple views for autostereoscopic displays in a single rendering pass. Our algorithm is based on the fact that vertices properties remain the same from different viewpoints. Taking advantage of the latest generation of GPUs including geometry shaders, we propose a method that significantly speeds up the rendering process by duplicating and transforming incoming primitives for a defined set of views. Our method involves very few modifications to be used with a standard stereo device.

The paper will be presented by

François de Sorbier

---

**Neural Networks for Arm Movement Prediction in CVEs**

Fred Stakem and Ghassan AlRegib (GaTech, USA)

Whether interacting with a Collaborative Virtual Environment, or CVE, locally or one networked across the Internet, any delay in the system can lead to a reduced sense of immersion. Input sensor delay and network delay are two common problems in CVE design that can be overcome with the application of prediction algorithms to the system. The purpose of this experiment was to assess the quality of feed forward back propagation neural networks in predicting natural avatar arm movement used in a CVE. In addition the experiment attempted to find the bounds for precise neural network prediction. The results show many different combinations of back propagation neural network topologies are capable of predicting up to 400 ms of human arm movements relatively accurately.

The paper will be presented by

Fred Stakem
3D-Model view characterization using equilibrium planes
Adrien Theetten, Tarik Filali Ansary and Jean-Philippe Vandeborre (LIFL, France)

We propose a new method for 3D-mesh model characteristic view selection. It consists in using the views that come from the equilibrium states of a 3D-model: they correspond to the horizontal plane on which an object is statically laying under the effect of gravity. The selected views are then very intuitive for the user. Indeed, to present a query, the user will take a photo or draw a sketch of the object on a table or on a floor, putting thus the object in a static mechanical equilibrium.

Consequently, our view selection method follows the same principles: finding all the equilibrium planes of an object and obtaining their relative 2D views. We present the experiments and results of our method on the Princeton 3D Shape Benchmark Database using a collection of 50 images (photos, sketches, etc.) as queries, showing the performance of our method in 3D retrieval from photos.

The paper will be presented by Jean-Philippe Vandeborre

Using Markov Random Fields and Algebraic Geometry to Extract 3D Symmetry Properties
Yunfeng Sui and Andrew Willis (UNC-Charlotte, USA)

In this paper, we present a new technique for solving the difficult problem of estimating the axis of symmetry for axially-symmetric surfaces. Accurate solutions to this problem are important in archaeology for systems that seek to reconstruct pottery vessels from measurements of their fragments. Our approach estimates quadratic surfaces at each measured surface point and uses a Markov Random Field superimposed on the measured surface mesh to estimate a collection of surface patches, each of which lies close to a single 3D quadratic surface. For each surface patch we estimate an quadratic implicit polynomial whose coefficients directly provide an estimate of the unknown axis location and orientation. Competing estimates of the global axis are combined using a Maximum Likelihood Estimation (MLE) framework that reflects the uncertainty present in the estimates computed from each surface patch. Our approach differs from past approaches by combining estimates derived from large surface regions that include many measurements instead of combining many local (often pointwise) estimates of the surface to determine the global estimate. Estimates from these large regions are more robust to noise and have sufficient data to generate statistics that accurately reflect the uncertainty in the computed estimates. As such, each estimate of the central axis is less susceptible to outliers and the overall axis estimate is significantly improved.

The paper will be presented by Yunfeng Sui

An efficient and memory-conserving implementation of multi-view stereo for wide-area reconstruction
Xenophon Zabulis (ICS-FORTH, Greece), Nikolaos Grammalidis (ITI-CERTH, Greece) and Georgios D. Floros (ICS-FORTH, Greece)

This paper deals with the automatic stereo reconstruction of wide-area scenes. Its particular goal is a computationally efficient method that can be performed on a personal computer, despite the large amount of data involved in the reconstruction of wide-area scenes. Robustness is considered in terms of the accuracy of the final reconstruction, as well as, in the context of simplifying the image acquisition process for the end-user.

The paper will be presented by Xenophon Zabulis
Fast and High Quality Fusion of Depth Maps
Christopher Zach (UNC-Chapel Hill, USA)

Reconstructing the 3D surface from a set of provided range images -- acquired by active or passive sensors -- is an important step to generate faithful virtual models of real objects or environments. Since several approaches for high-quality fusion of range images are already known, the runtime efficiency of the respective methods are of increased interest. In this paper we propose a highly efficient method for range image fusion resulting in very accurate 3D models. We employ a variational formulation for the surface reconstruction task. The global optimal solution can be found by gradient descent due to the convexity of the underlying energy functional. Further, the gradient descent procedure can be parallelized, and consequently accelerated by graphics processing units. The quality and runtime performance of the proposed method is demonstrated on well-known multi-view stereo benchmark datasets.

The paper will be presented by Christopher Zach

Efficient 3D Shape Acquisition and Registration Using Hybrid Scanning Data
Hongwei Zheng, Dietmar Saupe, Markus Roth, Andreas Boehler and Peter Opuchlik (University of Konstanz, Germany)

We consider efficient 3D shape acquisition and surface registration using dissimilar laser range scanners. In this paper, we exploit the fundamental 3D scanning trade-off between the coverage of the global shape structure and numerous surface patches to construct a hybrid laser scanning system provided that it can acquire both global and local shape information. The scanned low-resolution global shape data supplies the global shape structural prior for registering the high-resolution local 3D surface patches. Local surface patches can thus be optimally registered requiring less overlapping and thus reducing redundancy. To verify the feasibility of our hybrid 3D laser scanning system, we have implemented a prototype based on two laser range scanners, a hand-held one for the coarse global low-resolution model and a second stationary high-resolution line scanning system. This prototype system was evaluated for various real 3D models. Based on geometric data alone without using texture information, the results show that the proposed hybrid 3D scanning system outperforms previous approaches to the 3D acquisition and surface registration problem. The approach can be further extended and applied to other practical 3D shape applications.

The paper will be presented by Hongwei Zheng

Demo
19:00 to 22:00 - June, 18 (Wednesday)
(with the poster session)

A Range Scanner for Moving Objects
Huynh Quang Huy Viet (Saitama University, Japan) and Hiroshi Kawasaki (Saitama University, Japan), Ryo Furukawa (Hiroshima City University, Japan) and Ryusuke Sagawa, (Osaka University, Japan) and Yasushi Yagi (Osaka University, Japan)

We demonstrate a 3D scanning system consisting of a single camera and a single video projector, which can capture a dense 3D shape of a moving object, e.g. rotating fan, human body, etc. The system is based on our original ‘one-shot’ scanning technique which uses a single dense grid pattern with only two colors; therefore, the technique is robust in terms of image processing. Another advantage of the system is that no synchronization device between a CCD camera and a video projector is necessary for the system, and thus, an inexpensive system can be realized with off-the-shelf devices and a simple configuration.
Enhancing Textured Digital Elevation Models Using Photographs
Martin Schneider and Reinhard Klein (University of Bonn, Germany)

We present a method for enhancing the visual quality of existing digital elevation models textured with orthophotos, using a sparse set of unordered, high resolution photographs. After an initial manual selection of correspondence points, we automatically register the input photographs to the given terrain data set using robust image-based modeling techniques. To combine the geo-registered images on the terrain surface, we propose a compositing algorithm that ensures smooth transitions between the images while at the same time preserving the fine details. The resulting textures are inserted into the quadrtree representation of a terrain rendering engine to allow an efficient realtime visualization. We demonstrate our method on an HRSC terrain data set and a collection of high resolution photos of Turtmann valley in Switzerland.

The paper will be presented by Martin Schneider

A Three-tier Hierarchical Model for Capturing and Rendering of 3D Geometry and Appearance from 2D Images
Martin Jagersand, Neil Birkbeck and Dana Cobzas (U of A, Canada)

We propose a three-scale hierarchical representation of scenes and objects and show how this representation is suitable for both capture of models from images and efficient photo-realistic rendering. The model consists of: (1) a conventional triangulated geometry on the macro-scale; (2) a displacement map, introducing pixel-wise depth with respect to each planar model facet (triangle) on the meso level; (3) a photo-realistic micro-structure represented by an appearance basis spanning viewpoint variation in texture space. We implement a capture and rendering system for this model. Conventional Shape-From-Silhouette or Structure-From-Motion is used to capture the coarse macro geometry, variational shape and reflectance estimation for the meso-level, and texture basis optimization for the micro level. For efficiency the meso and micro level routines are both HW accelerated. Photo-realistic capture of complex scenes is thus possible in a few minutes using budget cameras and PCs, and rendering is real-time. Experimental results and videos show models from regular images of humans and objects.

The paper will be presented by Neil Birkbeck

Smooth and non-smooth wavelet basis for capturing and representing light
Cameron Upright, Dana Cobzas and Martin Jagersand (U of A, Canada)

Estimating light sources from images is an important and difficult problem in computer vision. High quality lighting is useful as input to other computer vision algorithms and in graphics rendering. For instance, photometric stereo and shape from shading requires known light and are limited to highly controlled laboratory setups. Accurate lighting is also helpful in augmented reality in order to consistently relight an artificially introduced object. Algorithms using individual point lights are useful for simple lighting setups, but for complex illumination a basis representation is needed. We propose a lighting model using Daubechies wavelets and a method for recovering light from cast shadows and specular highlights in images. We assume that the geometry is known for part of the scene. We tested our method for difficult cases of both uniform and textured objects and under complex geometry and light conditions, and show good results using the proposed Daubechies basis on both synthetic and real datasets.

The paper will be presented by Cameron Upright
An Energy Formulation for Establishing the Correspondence Used in Projector Calibration

Marc-Antoine Drouin (CNRC, Canada), Guy Godin (CNRC, Canada) and Sébastien Roy (UDEM, Canada)

Projector-camera systems are used in many applications that need geometric calibration of the projectors. As is the case for cameras, calibration of a projector requires the establishment of the correspondence between known features of a calibration target and the projector pixels. However, a fundamental difference is that the method must rely on external sensors (i.e. cameras) for its calibration. This introduces additional sources of error that affect the accuracy of the projector calibration. As a solution, we propose a multi-camera energy-based approach for establishing the correspondence between the projector and a calibration target. Performing projector calibration using the correspondence obtained by our method allows a significant improvement in accuracy. We demonstrate the validity of our approach using off-the-shelf and low cost equipment, and validate the results using a laser tracker.

The paper will be presented by Marc-Antoine Drouin

Integration of a Time-of-Flight Camera into a Mixed Reality System for Handling Dynamic Scenes, Moving Viewpoints and Occlusions in Real-Time

Bogumil Bartczak, Ingo Schiller, Christian Beder and Reinhard Koch (CAU-Kiel, Germany)

A novel approach to mixed reality applications is discussed. The key characteristics of the presented system are the use of an automatically generated static environment model and a time-of-flight camera device. The combination of both allows the correct handling of mutual occlusion between real and virtual content on the fly, which is not possible with the currently applied approaches. Typically expensive studio setups with complex camera tracking installations and multi-camera approaches in combination with chroma-keying facilities are used. The system is rather inexpensive, compact, mobile, flexible and provides convenient calibration procedures. The use of a background model not only eliminates the need for chroma-keying in mixed reality production, it moreover supports planing and alignment of virtual content. Based on depth information the system is generating appropriate depth maps in real-time, making the approach suitable for 3D-TV productions. The presented paper discusses all key elements of mixed-reality applications based on this approach. This includes camera pose tracking, correct real-time handling between interacting virtual and real content and the fast environment model building.

The paper will be presented by Bogumil Bartczak

Proxy-Based Compression of 2-1/2 D Structure of Dynamic Events for Tele-immersive Systems

Pooja Verlani and P.J. Narayanan (IIIT-Hyderabad, India)

The 2-1/2D geometric structure and photometric appearance of dynamic scenes find applications in 3D tele-immersive systems. The captured "depth movies" contain aligned sequences of depth maps and textures and are often streamed to a distant location for immersive viewing. Depth movies are heavy and need to be compressed. We present a scheme to compress depth movies of human actors using a parametric proxy model for the underlying action. A generic articulated human model is used as the proxy and its joint angles are the parameters for each time instant. The proxy represents a common prediction of the scene structure. The residue between the captured depth map and the proxy depth map is used to represent the scene. This exploits the spatial coherence between multiple depth movies. Temporal coherence is exploited by encoding the difference between successive frames of residues. Intra-frame coded frames and difference-coded frames provide random access and high compression. Results on synthetic and real actions demonstrate an analysis of the compression ratio and resulting quality of the decoded scene.

The paper will be presented by Pooja Verlani
MRF Stereo with Statistical Parameter Estimation
Shafik Huq, Andreas Koschan, Besma Abidi and Mongi Abidi (UTK, USA)

A Markov Random Field (MRF) based local stereo matching algorithm that estimates parameters automatically from statistics is proposed. For an iterative optimization, cost functions working on local support neighborhood are developed. Data model parameters are pre-estimated from one of the stereo images by applying a noise equivalence hypothesis. The smoothness model parameters are estimated with maximum likelihood (ML) applying disparity gradient constraint and 3*sigma confidence boundary. The confidence boundary also defines the parameters for handling discontinuities in data and smoothness. Additionally, homogeneous points are included into the support neighborhood to achieve high matching rate along surface borders. Finally, a pair of cost functions is modeled to match the images symmetrically for improved matching. Experiments on ground truth datasets show that among the existing algorithms with statistical estimation of the parameters, the proposed algorithm delivers the highest matching rate.

The paper will be presented by
Andreas Koschan

Stereo matching using interior point methods
Arvind Bhusnurmath and Camillo J. Taylor (UPenn, USA)

This paper describes an approach to reformulating the stereo matching problem as a large scale Linear Program. The approach proceeds by approximating the match cost function associated with each pixel with a piecewise linear convex function. Regularization terms related to the first and second derivative of the disparity field are also captured with piecewise linear penalty terms. The resulting large scale linear program can be tackled using interior point methods and the associated Newton steps involve Hessian matrices that reflect the structure of the underlying pixel grid. The proposed scheme effectively exploits this structure to efficiently solve the large scale global optimization problem.

The paper will be presented by
Arvind Bhusnurmath

Marker-less motion capture of skinned models in a four camera set-up using optical flow and silhouettes
Luca Ballan and Guido Maria Cortelazzo (UniPD, Italy)

We explore a new approach to marker-less motion tracking of a priori known skinned meshes using both optical flow and silhouette information. We present a formulation which considers in a unified way both these two kinds of information and accounts for the non-rigid deformations of the object skin modeling them using the Skeletal Subspace Deformation (SSD). We then demonstrate the effectiveness of our technique showing its performance in a four camera set-up tracking a subject modeled by a skeleton with 46 degrees of freedom.

The paper will be presented by
Luca Ballan
**Invited Session**

10:30 to 12:30 - June, 20 (Friday)

**Challenges in wide-area structure-from-motion**

Marc Pollefeys (ETH-Zurich, Switzerland and UNC-Chapel Hill, USA)

In this talk, I will present work on wide-area structure-from-motion (SfM) such as city-wide 3D reconstructions from millions of video frames. While in recent years a lot of progress has been made in the area of SfM and multi-view stereo reconstruction, wide-area 3D reconstructions and mapping lead to interesting new research challenges. It is for example important to use algorithms that are efficient as typically millions of frames have to be processed. In this context we will present algorithms which exploit the tremendous computational power of recent graphic processing units (GPU) to achieve real-time performance. Another challenge consists of avoiding unbounded accumulation of errors. For this it is important to close loops when the camera path crosses itself (e.g. at street intersections). We introduce viewpoint-invariant-patches (VIP) to enable robust and efficient matching over widely varying viewpoints (e.g. orthogonal crossings). Our approach is illustrated with 3D reconstructions of Chapel Hill (where the last edition of 3DPVT was held).

Marc Pollefeys

**Photosynth and Beyond**

Drew Steedly (Microsoft Live Labs, USA)

Traditional photo browsing tools allow you browse through 2D pages of thumbnails. Photosynth lets you to browse photo collections in 3D space. This makes it easy to answer questions like “What is to the right of this photo?” and “Is there a more detailed photo of this part of the scene?”. The Photosynth viewing experience relies on first automatically reconstructing the camera positions and a sparse point cloud. In this talk, I will discuss some recent enhancements to the viewing experience as well as a tool that allows users to interactively build textured 3D models from Photosynths.

Drew Steedly

**Taking Google Maps to Street Level**

Jana Kosecka (GMU and Google, USA)

I will describe Google's Streetview feature from the conception of the idea to initial capture experiments, challenges encountered along and present and future directions. I will talk about what it takes to increase the quantity of coverage while maintaining and improving the quality of one of the few application where the number of images can be measured in miles.

Jana Kosecka
3D modeling of large scale environments is of importance in many applications such as city planning, training and simulations, architectural studies, gaming and entertainment, and emergency services. In this talk, we describe an approach to fast, automated, 3D modeling of large scale environments using airborne data only. In contrast with ground based modeling which entails driving on every street of a city, airborne data acquisition can be significantly faster, and hence can scale to much larger areas.

Our basic approach is to construct the 3D geometry using airborne LiDAR data obtained by an airplane, and to texture map this model using aerial imagery from a helicopter equipped with inexpensive inertial measurement units (IMU). At the core of our approach lies an automated algorithm for texture mapping oblique aerial images onto a 3D model generated from airborne LiDAR data. Our proposed texture mapping algorithm consists of two steps. In the first step, we combine vanishing points and global positioning system aided inertial system readings to roughly estimate the extrinsic parameters of a calibrated camera. In the second step, we refine the coarse estimate of the first step by applying a series of processing steps. Specifically, we extract 2D orthogonal corners (2DOCs) corresponding to orthogonal 3D structural corners as features from both images and the untextured 3D LiDAR model. The correspondence between an image and the 3D model is then performed using Hough transform and generalized M-estimator sample consensus. The resulting 2DOC matches are used in Lowe’s algorithm to refine camera parameters obtained earlier. Our system achieves 91% correct pose recovery rate for 90 images over the downtown Berkeley area, and overall 61% accuracy rate for 358 images over the residential, downtown and campus portions of the city of Berkeley.

Reconstruction

14:00 to 16:00 - June, 20 (Friday)
Session chair: Marc Pollefeys

3D Object Reconstruction with Heterogeneous Sensor Data

Li Guan (UNC-Chapel Hill, USA and ETH-Zurich, Switzerland), Jean-Sébastien Franco (LaBRI, France) and Marc Pollefeys (ETH-Zurich, Switzerland and UNC-Chapel Hill, USA)

In this paper, we reconstruct 3D objects with a heterogeneous sensor network of Range Imaging (RIM) sensors and high-res camcorders. With this setup, we first carry out simple but effective depth calibration for the RIM cameras. We then combine the camcorder silhouette cues and RIM camera depth information, for the reconstruction. Our main contribution is the proposal of a sensor fusion framework so that the computation is general, simple and scalable. Although we only discuss the camcorders and RIM cameras in this paper, the proposed framework can be applied to any type of vision sensors. It uses a space occupancy grid as a probabilistic 3D representation of scene contents. After defining sensing models for each type of sensors, the reconstruction is simply a Bayesian inference problem, and can be solved robustly. The experiments show that the recover full 3D closed shapes substantially improved the quality of the noisy RIM sensor measurement.

The paper will be presented by Li Guan
L-Tangent Norm: A Low Computational Cost Criterion for Choosing Regularization Weights and its Use for Range Surface Reconstruction

Florent Brunet (CAMPAR, Germany and LASMEA / LAIC, France), Adrien Bartoli (LASMEA, France), Rémy Malgouyres (LAIC, France) and Nassir Navab (CAMPAR, Germany)

We are interested in fitting a surface model such as a tensor-product spline to range image data. This is commonly done by finding control points which minimize a compound cost including the goodness of fit and a regularizer, balanced by a regularization parameter. Many approaches choose this parameter as the minimizer of, for example, the cross-validation score or the L-curve criterion. Most of these criteria are expensive to compute and difficult to minimize. We propose a novel criterion, the L-tangent norm, which overcomes these drawbacks. It gives sensible results with a much lower computational cost. This new criterion has been successfully tested with synthetic and real range image data.

The paper will be presented by Florent Brunet

GPU Accelerating Speeded-Up Robust Features

Timothy Terriberry, Lindley French and John Helmsen (ArgonST Inc., USA)

Many computer vision tasks require interest point detection and description, such as real-time visual navigation. We present a GPU implementation of the recently proposed Speeded-Up Robust Feature extractor, currently the state of the art for this task. Robust feature descriptors can give vast improvements in the quality and speed of subsequent steps, but require intensive computation up front that is well-suited to inexpensive graphics hardware. We describe the algorithm's translation to the GPU in detail, with several novel optimizations, including a new method of computing multi-dimensional parallel prefix sums. It operates at over 30 Hz at HD resolutions with thousands of features, and in excess of 70 Hz at SD resolutions.

The paper will be presented by Timothy B. Terriberry

Single-Step Planar Surface Extraction from Range Images

Shanmugalingam Suganthan, Sonya Coleman and Bryan Scotney (University of Ulster, UK)

The volume of raw range image data that is required to represent just a single scene can be extensive; hence direct interpretation of range images can incur a very high computational cost. Range image segmentation and feature extraction have been identified as mechanisms to produce more compact scene representations and hence enable less costly scene interpretation for applications such as object recognition and robot navigation. We present a new approach to multi-scale edge and planar surface detection in range images that can be used directly with any range data, regardless of whether the data have regular or irregular spatial distribution. The approach is evaluated with respect to accuracy of both edge location and planar surface representation. The contribution of our approach is that only a single application of our operator is necessary to both extract object edges and determine representations of the corresponding planar object surfaces.

The paper will be presented by Shanmugalingam Suganthan
The conference will be held at the Klaus Advanced Computing Building (KACB, room 1443), located in the center of campus, at 266 Ferst Drive, Atlanta, GA 30332. The Tech Trolley connects the KACB, Georgia Tech Hotel, and the Midtown station of Atlanta’s public transit system, MARTA. The Tech Trolley will run approximately every 12 minutes.

Lunch and Dinner Breaks
Atlanta has a stellar variety in restaurants, affordable, expensive and everywhere in between. We list a few of our favorites for lunch and dinner. Please also consult the restaurant pages of http://www.accessatlanta.com/restaurants for more tips and reviews.

Lunch & Dinner
- Baraonda (designer pizza). 710 Peachtree Street. 404-879-9962 $$
- Eno Restaurant and Wine Bar (Mediterranean). 800 Peachtree Street. 404-685-3191 $$$
- F2O (salads) 674 Myrtle Street $
- 5th Street Ribs & Blues (BBQ). 86 5th Street, N.W. 404-249-8808 $
- Fune Sushi Bar. 860 Peachtree Street at 7th. 404-541-9322 $$
- Gladys Knight (Southern/soul food). 529 Peachtree Street. 404-874-9393 $$
- Little Azio’s (pizza). 903 Peachtree Street. 404-876-7711 $
- Pleasant Peasant. (French Bistro) 555 Peachtree Street, N.E. 404-874-3223 $$$
- The Globe (eclectic bistro). 75 5th Street N.E. 404-541-1487 $
- The Spotted Dog (burgers/salads) 30 North Avenue. N.W. 404-347-7337 $
- Tin Drum (Thai/noodles). 88 5th Street. 404-881-1368 $
- Toast (bistro). 817 West Peachtree Street. 404-815-9243 $$
- The Vortex (brewpub). 878 Peachtree. 404-875-1667 $

Dinner
- Nan Thai. 1350 Spring Street. 404-870-9933 $$$
- Fuego Cafe and Tapas Bar. 1136 Crescent Avenue. 404-389-0660 $$
- Nikiemoto’s (sushi/fusion). 990 Piedmont Avenue. 404-892-4111 $$
- Las Palermas (Cuban). 368 5th Street N.E. 404-872-0846 $$
- MF Sushi Bar. 265 Ponce de Leon. 404-541-9997 $$
- Trois (French). 1180 Peachtree Street. 404-815-3337 $$
Vegetarian
• Cafe Sunflower. 2140 Peachtree Rd. N.W. 404-352-8859 $
• Dressed. 950 W. Peachtree St. 404-347-3434 $
• Green Sprout. 1520 Piedmont Avenue. 404-874-5336 $
• Olive Bistro. 650 Ponce de Leon (Greek/Mediterranean). 404-874-5336 $
• R. Thomas Deluxe Grill. 1812 Peachtree Rd. 404-881-0246 $

Hemphill Area
• Rocky Mountain Pizza. 1005 Hemphill Ave N.W. 404-876-8600
• City Café Diner. 525 10th St. N.W. 404-724-0407
• Lucky Buddha (vegetarian Asian). 525 10th St. 404-885-1518
• Firehouse Subs. 525 10th St. N.W. 404-347-9912
• Donuts. 525 10th St. N.W.

Student Center
• Pizza Hut, Chick-fil-A, Far East Fusion, Salad Bar, The Chef’s Line, Ferst Place (Gourmet Lunch), Einstein’s Bagels

Breakfast
• Starbucks. 5th Street (next to Georgia Tech Bookstore)
• Donuts. 525 10th St. N.W. (Hemphill Area)
• Student Center

About Atlanta
Overview: Atlanta is the capital of the state of Georgia, and the central city of the ninth most populous metropolitan area in the United States, with a population of 5,478,667. In the last decade, the Atlanta area added over 1,150,000 residents and is recognized as one of the driving forces of the “New South”. During the Civil Rights Movement, Atlanta stood apart from Southern cities that supported segregation, touting itself as the “City Too Busy to Hate.” The city’s progressive civil rights record made it increasingly popular as a relocation destination for African Americans, who became the dominant political force in the city; since 1974, all of the mayors of Atlanta have been African-American. The city is divided in three parts: Downtown, the traditional commercial center of Atlanta; Midtown, up and coming commercial and residential area; and Buckhead, the financial district.

Museums: Atlanta hosts a variety of museums. Prominent among them are sites honoring Atlanta’s participation in the civil rights movement. Reverend Dr. Martin Luther King, Jr. was born in the city, and his boyhood home on Auburn Avenue in the Sweet Auburn district is preserved as the Martin Luther King, Jr. National Historic Site. Other history museums and attractions include the Atlanta History Center; the Atlanta Cyclorama and Civil War Museum; the Carter Center and Presidential Library; and the Margaret Mitchell House and Museum. The arts are represented by several theaters and museums, including the Fox Theatre. The High Museum of Art is the city’s major fine/visual arts venue, with a significant permanent collection and an assortment of traveling exhibitions. Atlanta features the world’s largest aquarium, the Georgia Aquarium, which features over 100,000 specimens in tanks holding approximately eight million gallons of water. A unique museum is the World of Coca-Cola featuring the history of the world famous soft drink brand and its well-known advertising. Next door is the CNN building, where tours are organized to show the inner workings of a cable news station.

Parks: The heart of the city is Piedmont Park, with Ultimate Frisbee meet-ups on Saturdays around 10:30 am, and the Atlanta Botanical Gardens. Just east of the city, Stone Mountain is the largest piece of exposed granite in the world. On its face are giant carvings of Jefferson Davis, Robert E. Lee, and Stonewall Jackson. It is also the site of impressive laser shows in the summer.