10371 Abstracts Collection
Dynamic Maps
— Dagstuhl Seminar —

Claus Brenner[1], Wolfram Burgard[2], Marc Pollefeys[3] and Christoph Stiller[4]

[1] Leibniz Universität Hannover, DE
claus.brenner@ikg.uni-hannover.de
[2] Universität Freiburg, DE
burgard@informatik.uni-freiburg.de
[3] ETH Zürich, CH
marc.pollefeys@inf.ethz.ch
[4] KIT Karlsruhe, DE
stiller@mrt.uka.de

Abstract. From September 12th to 17th, 2010, the Dagstuhl Seminar 10371 "Dynamic Maps " was held in Schloss Dagstuhl – Leibniz Center for Informatics. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

Keywords. Dynamic maps, computer vision, photogrammetry, robotics, computer graphics, geoinformatics, driver assistance

10371 Executive Summary – Dynamic Maps

In recent years, the advent of car navigation systems has laid the ground for an entirely new industry sector, consisting of map producers, car/ personal/ smart phone navigation manufacturers, and service providers. It has probably gone unnoticed that navigation systems mark a major change in the way we use maps. Partially, they are still just a replacement for traditional maps, providing a means to store and visualize a representation of the environment. In contrast to the traditional use of maps, however, navigation systems perform computations using the map's data structures, such as shortest route, map matching, and route guidance. That is, from an abstract point of view, part of the map is made for machine use only - the user has no direct access to it but rather is only presented the outcome of the computations.

Keywords: Dynamic maps, spatial data infrastructure, SLAM, map matching, autonomous navigation, spatial cognition, data fusion, information retrieval, intelligent vehicles, 3D scene perception, scene understanding, 3D reconstruction

Joint work of: Brenner, Claus; Burgard, Wolfram; Pollefeys, Marc; Stiller, Christoph

Full Paper:
http://drops.dagstuhl.de/opus/volltexte/2010/2951

Towards Semantic Maps for Domestic Service Robots

Sven Behnke (Universität Bonn, DE)

In the talk, I cover three different aspects of spatial representations needed for domestic service robots:

- maps of domestic environments for localization and mapping,
- representations of movable objects for manipulation, and
- representation of persons in the environment.

The talk is illustrated using examples from our domestic service robot Dynamaid, which successfully competes at RoboCup@Home competitions. The examples include continuous person awareness, gesture recognition, object detection, object recognition, analysis of table-top scenes, and 3D obstacle avoidance.

Humanoid Navigation in Complex Indoor Environments

Maren Bennewitz (Universität Freiburg, DE)

Humanoid robots have become a popular research tool in recent years. The motivation behind working with humanoids is that these robots are capable of human-like navigation, i.e., they can climb stairs or move away objects blocking their way. Our goal is to enable humanoid robots to autonomously navigate in complex indoor environments consisting of multiple levels connected by staircases and containing movable and articulated objects. In this talk, I give an overview about our work on 3D environment modeling, localization, path planning, and navigation with humanoid robots.

Non-Rigid Registration and Rectification of 3D Laser Scans

Dorit Borrmann (Jacobs University Bremen, DE)

Three dimensional point clouds acquired by range scanners often do not represent the environment precisely due to noise and errors in the acquisition process. These latter systematical errors manifest as deformations of different kinds in the 3D range image. We have developed a novel approach to correct deformations by an analysis of the structures present in the environment and correcting them by non-rigid transformations. The resulting algorithms are used for creating high-accuracy 3D indoor maps. When a scan is acquired during movement of the mobile robot new errors are introduced into the system. We want to investigate how these errors can be detected and rectified.

Related references

[1] Elseberg, J., Borrmann, D., Lingemann, K., and Nüchter, A., Non-Rigid Registration and Rectification of 3D Laser Scans, IEEE/RSJ International Conference on Intelligent Robots and System (IROS 2010), Taipei, Taiwan, October, 2010.

Keywords: Robotic mapping, SLAM, laser scanning, rectification of laser scans, non-rigid registration, thin plate splines

Full Paper:
http://plum.eecs.jacobs-university.de/download/iros2010_1.pdf

Joint work of: Borrmann, Dorit; Elseberg, Jan; Lingemann, Kai; Nüchter, Andreas

Dynamic Maps - a Geodesist's Perspective

Claus Brenner (Leibniz Universität Hannover, DE)

This talk gives an overview, covering aspects of data acquisition, mapping and information extraction from raw data such as laser scans, from a photogrammetry/cartography perspective. There are also some remarks regarding the typical modelling/ontology for current maps, as found in ATKIS and GDF (car navigation) maps and typical problems and operations found in map processing, such as integration of different maps, cartographic generalization, typification, and aggregation. Finally, todays view of 'dynamics' is outlined, with an example of change detection for map updating.

Mapping with Mobile Robots in Dynamic Environments

Wolfram Burgard (Universität Freiburg, DE)

The problem of generating maps with mobile robots has received considerable attention over the past years. Most of the techniques developed so far, however, have been designed for situations in which the environment is static during the mapping process. Dynamic objects in the environment can lead to serious errors in the resulting maps such as spurious objects or misalignments due to localization errors. In this talk we consider the problem of creating maps with mobile robots in dynamic environments. We present different probabilistic approaches for estimating dynamic objects, for simultaneous mapping and localization augmented by a probabilistic technique for identify spurious measurements, for identifying semi-static structures, and for learning sub-maps of dynamic areas. In several experiments we demonstrate the properties and limitations of the individual algorithms.

Keywords: Robot Mapping, Dynamic Environments

Computer Vision for Dynamic Maps

Jan-Michael Frahm (University of North Carolina - Chapel Hill, US)

In this talk I review the recent work towards dynamic maps in the area of computer vision. In particular I review the state of the art in photo collection modeling, real-time reconstruction from video, fast location recognition, change detection, and modeling dynamic objects in the scene. Additionally, I will discuss the challenges to be addressed in the future.

Local Environment Perception and Global Maps – a Car Manufacturer's View

Uwe Franke (Daimler AG, DE)

The next generation of driver assistance systems will use powerful local environment perception in order to avoid serious accidents. Their situation analysis is based on local maps derived from the vehicle's sensors. There is no doubt that those systems could benefit from rich global map information. Scientific as well as commercial projects are on the way to deliver this information.

In my talk I will first sketch the state-of-the-art in real-time spatio-temporal computer vision and local modelling. Secondly, I will point out the benefits driver assistance would gain from precise maps and how – from the current point of view – we would use this information in future applications.

Keywords: Driver assistence, computer vision, local maps

Visual localization and its need for 3D maps and semantics

Friedrich Fraundorfer (ETH Zürich, CH)

One specific field of computer vision is the problem of visual localization, to compute the answer to the question "Where am I" based on image content.

In this talk I will explain a method in which visual localization can be seen as image search in a database that contains image data annotated with location information. I will discuss details of which visual information should be stored in the database and how it can be searched efficiently. And I will also discuss the role of 3D information in form of 3D models and how this 3D data can be obtained.

I will demonstrate several different applications of visual localization, i.e., location recognition for mobile devices, loop closing for mobile robots, image search in large databases, robust structure-from-motion and interactive virtual tours.

Towards building a map of a collapsed building with a range camera on an actuated endoscope

Udo Frese (Universität Bremen, DE)

If a building collapsed typically rescue dogs search for victims. But rescue dogs only communicate that "someone is down there" not precisely, where the victim is. The victim must then be localized, a task poorly supported by current technology. Our vision is to mount a range-camera on an acuated endoscope to search the rubble pile. A 3D mapping system should support the operator, because endoscope images are very hard to grasp spatially. For this task a range-camera is an ideal sensor, however it is known for involving many peculiar effects in the data.

Our open question is: Can we utilize the redundancy in viewing every point of the scene many times to get the best from that difficult data? We present this open question as well as initial thoughts on phrasing it as a probabilistic problem.

Keywords: SLAM, range-camera, rescue robotics, USAR, probabilistic mapping

Visual Attention for Mobile Systems

Simone Frintrop (Universität Bonn, DE)

Visual attention is a concept of human perception that directs the gaze to scene regions of potential interest. This is especially important if an autonomous agent has to navigate and act in a complex world. I show an application of scene recognition and visual SLAM based on salient landmarks. These landmarks are highly repeatable and discriminative and allow a sparse landmark representation.

g2o: A General Framework for Graph Optimization

Giorgio Grisetti (Universität Freiburg, DE)

Many popular problems in robotics and computer vision including various types of simultaneous localization and mapping (SLAM) or bundle adjustment (BA) can be phrased as least squares optimization of an error function that can be represented by a graph. This talk describes the general structure of such problems and presents g2o, an open-source C++ framework for optimizing graph-based nonlinear error functions. Our system has been designed to be easily extensible to a wide range of problems and a new problem typically can be specified in a few lines of code. The current implementation provides solutions to several variants of SLAM and BA. We provide evaluations on a wide range of real-world and simulated datasets. The results demonstrate that while being general g2o offers a performance comparable to implementations of state-of-the-art approaches for the specific problems.

Keywords: SLAM, least squares, graph optimization, bundle adjustment

Joint work of: Grisetti, Giorgio; Kuemmerle, Rainer; Strasdat, Hauke; Konolige, Kurt; Burgard, Wolfram

Localization in a map of landmarks: an approach by point pattern matching

Jan-Henrik Haunert (Universität Würzburg, DE)

With the rapid evolution of advanced driver assistance systems, the need for accurate and robust vehicle positioning methods increases. The Global Positioning System (GPS), which today is widely used for navigation, is not accurate enough, for example, to decide on which road lane a vehicle is moving. Moreover, a GPS-based positioning may fail due to occlusions, especially in urban areas. Therefore, other systems are to be applied instead of or in combination with GPS.

A promising approach to vehicle localization is to match observations of vehicle-mounted sensors (for example, laser scanners or cameras) with a reference dataset of the environment – indoor localization of robots is often done this way. When moving outdoors, however, the localization problem becomes a lot more complicated. For example, the placement of artificial targets at prominent positions, which is often done to ease indoor localization, becomes infeasible when the public outdoor environment is to be covered. The time is now to decide about which features to include in digital maps for automatic outdoor localization, how to keep them up to date, and – this will be the focus of my talk – how to solve the actual localization problem using such maps.

In this talk, I will present a new algorithm for vehicle localization by point-pattern matching. This algorithm first triangulates landmarks observed from the vehicle. Then, it matches the triangles in the triangulation with triangles of landmarks in a reference database based on their geometric similarity. The algorithm handles certain constraints that avoid conflicting triangle matches to become selected and guarantees an optimal solution with respect to a geometric quality measure.

In order to test this method, a reference database of pole-like landmarks was set up using a street mapper system. Furthermore, a vehicle was equipped with a low-cost laser scanner. By using the developed matching algorithm, it is possible to localize the vehicle accurately and reliably.

Keywords: Vehicle localization, map matching, point pattern matching, optimization

Autonomous Driving on the Way to Real-world Traffic

Soeren Kammel (BOSCH Research Center - Palo Alto, US)

Autonomous vehicle navigation made big progress through the DARPA Urban Challenge competition in 2007. Though demanding, the "toy" scenario in this contest - urban like driving in a mock up city wit traffic faked by stunt drivers - still was far from resembling real-world traffic. In this talk, the latest progress of Stanford's autonomous car Junior is presented in order to make this robot fit for driving safely among human drivers between Stanford and San Jose:

- Detection of traffic light states
- Jerk-optimal driving, reactive obstacle avoidance and merging into moving traffic
- Classification of laser range finder data with the purpose of pedestrian, bicyclist and vehicle detection and prediction

Keywords: Trajectory generation, classification, lidar, traffic light state detection

Geometry-aligned Image Feature Representations for Matching and Recognition

Kevin Koeser (ETH Zürich, CH)

When seeking correspondences between images, e.g., for localization or mapping, one has to deal with the variance in the image data that is caused by viewpoint and perspective. In this presentation methods are discussed to become (quasi-) invariant with respect to these nuisances. As an example, a system is presented for mobile phone based vision-only location recognition in urban scenarios, where street-level image data as well as rough 3D building models are exploited for 3D rotationally invariant representation of the facades of a city. Finally, dynamic aspects of this representation are discussed, e.g., regarding temporal updates of the model.

Towards Life Long Dynamic Maps

Jana Kosecka (George Mason Univ. - Fairfax, US)

Motivated by the problems of creating and maintaining life long dynamic maps for robotic agents I will briefly discuss three different representations of environments for and their potential for handling dynamic changes. These include topological mapping with discriminative feature selection, semantic mapping and metric 3D maps. These types of models incorporate both appearance based and

geometric cues and we have used them successfully for capturing indoors and outdoors environments and localization.

The issues related to the nature and suitability of these respresentations for increasingly more complex tasks, scalability and on-line updating will also be discussed.

Related references

[1] Micusik, B. and Kosecka, J., Piecewise Planar City Modeling from Street View Panoramic Sequences, IEEE CVPR 2009.
[2] Micusik, B. and Kosecka, J., Semantic Segmentation of Street Scenes by Superpixel Co-Occurrence and 3D Geometry, IEEE Workshop on Video-Oriented Object and Event Classification, ICCV 2009.
[3] Li, F. and Kosecka, J., Probabilistic Location Recognition Using Reduced Feature Set, IEEE International Conference on Robotics and Automation, Orlando, 2006.
[4] Zhang, W. and Kosecka, J., Image Based Localization in Urban Environments, International Symposium on 3D Data Processing, Visualization and Transmission, 3DPVT 2006.

Keywords: Dynamic maps

Full Paper:
http://cs.gmu.edu/∼kosecka/Publications.html

Towards a Semantic Interpretation of Dynamic Scenes

Bastian Leibe (RWTH Aachen, DE)

A main challenge in the creation of dynamic maps is the extraction of semantic information from the underlying image and video data. Such information is needed in order to separate transient from permanent structures or to remove individual object categories (such as pedestrians and cars) from the reconstructed data. In this talk, I will present examples from our recent work that demonstrate how such semantic information can be obtained on the object level from object categorization, on the sub-object (part) level from metadata transfer, and on the level of dynamics of moving objects from multi-person tracking. The unifying scenario for those components is a mobile vision system for recognizing and tracking other traffic participants from a moving vehicle. I will finish by discussing how semantic map information could in the future in turn be used to improve the performance of multi-person tracking systems.

Making Dynamic Maps Usable for Humans

Volker Paelke (IG - Castelldefels (Barcelona), ES)

Starting with a review of the classic cartographic communication model and the associated production pipeline the talk identifies the potential changes when moving from static paper maps to interactive real-time "geo-systems" that could constitute a dynamic map usable for human users. Different aspects ranging from data acquisition over processing to presentation and interaction are discussed to identify possible overlaps with other views of "dynamic maps", the associated techniques and algorithms and to identify possible areas of collaboration.

Keywords: Dynamic Maps, Visualization, Mixed Reality, Interaction, Humans, HCI, UI

1-Point RANSAC Structure from Motion and Absolute Scale Estimation from a Single Vehicle Mounted Camera by Exploiting Motion Nonholonomic Constraints

Davide Scaramuzza (ETH Zürich, CH)

The first biggest problem in visual motion estimation is data association; matched points contain many outliers that must be detected and removed for the motion to be accurately estimated. In the last few years, a very established method for removing outliers has been the "5-point RANSAC" algorithm which needs a minimum of 5 point correspondences to estimate the model hypotheses. Because of this, however, it can require up to thousand iterations to find a set of points free of outliers. In this talk, I will show that by exploiting the non-holonomic constraints of wheeled vehicles (e.g. cars, bikes, mobile robots) it is possible to use a restrictive motion model which allows us to parameterize the motion with only 1 point correspondence. Using a single feature correspondence for motion estimation is the lowest model parameterization possible and results in the most efficient algorithm for removing outliers: 1-point RANSAC.

The second problem in monocular visual odometry is the estimation of the absolute scale. I will show that vehicle non-holonomic constraints make it also possible to estimate the absolute scale completely automatically whenever the vehicle turns.

In this talk, I will give a mathematical derivation and provide experimental results on both simulated and real data over a large image dataset collected during a 25 km path.

Related references

[1] Scaramuzza, D., Fraundorfer, F., and Siegwart, R., Real-Time Monocular Visual Odometry for On-Road Vehicles with 1-Point RANSAC, IEEE International Conference on Robotics and Automation (ICRA 2009), Kobe, Japan, 16 May, 2009.

[2] Scaramuzza, D., Fraundorfer, F., Pollefeys, M., and Siegwart, R., Absolute Scale in Structure from Motion from a Single Vehicle Mounted Camera by Exploiting Nonholonomic Constraints, IEEE International Conference on Computer Vision (ICCV 2009), Kyoto, October, 2009.

Keywords: Structure from motion, visual odometry, SLAM, RANSAC, motion constraints

Full Paper:
http://drops.dagstuhl.de/opus/volltexte/2010/2950

Full Paper:
http://robotics.ethz.ch/~scaramuzza/Davide_Scaramuzza_files/Research/research.htm

Some challenges of dynamic cartography - an informal enquiry

Konrad Schindler (ETH Zürich, CH)

In the talk the creation and maintenance of dynamic maps is explored from the viewpoint of cartography. Additional challenges due to the dynamic nature of a map are identified in three areas: map revision in the presence of dynamically changing map content requires online updates and new data sources to detect and model changes of the topographic data; map generalization needs to be extended to include the consistent spatial generalization of dynamic content, and possibly also the generalization of temporal sequences; and map use requires the capability to predict the state of the dynamic world into the future, since maps typically serve as a basis for decisions about future action.

Incremental update of multi-scale and multi-temporal dynamic maps

Monika Sester (Leibniz Universität Hannover, DE)

Human perception is used to multi-scale representation of our environment. This is reflected in digital (+analog) maps as well by providing different scale-dependent representations. These different representations are thus generated by generalizing a highly detailed data set - or stem from observations made at smaller scales. In order to integrate incrementally acquired data from any sensor at any scale, the multi-scale property of objects has to be taken into account. Furthermore, also the time-dependent behavior of objects has to be known, in order to understand temporally varying objects.

Examples were shown how this knowledge can be used to verify newly acquired data, as well as an approach for cooperative information integration and refinement.

Keywords: Incremental update and refinement, multi-scale data

Recent work in image-based rendering from unstructured image collections and remaining challenges

Sudipta Sinha (Microsoft Research - Redmond, US)

I review recent progress in image-based rendering from unstructured images and discuss some remaining challenges. First, I describe our recent work on developing a multi-stage linear approach for structure from motion initialization and show that it is accurate but also much faster than the popularly used sequential approach. I then describe recent work on a piecewise planar stereo approach to compute compact piecewise planar 3d depth maps which are then used for view-interpolation. Finally I talk about some recent work where we perform image-based navigation of a scene from thousands of images but without performing a full global scene reconstruction.

Keywords: Structure from motion, image-based rendering, stereo

Full Paper:
http://research.microsoft.com/en-us/um/redmond/groups/ivm/LinearSfm/

PhotoCity: Collaborative 3D Reconstruction using Online Games

Noah Snavely (Cornell University, US)

There are many, massive sources of image data available on the web, including tourist photos on Flickr and other photo-sharing sites, and Google Street View and satellite imagery. However, there is still a largely untapped possible source of images: millions of people walking around with digital cameras and cell phones. If we could only encourage such people to take useful new photos, we would have access to a very large-scale distributed camera with views on all of the world's cities. This talk will try to encourage discussion about how this "camera" can be used, and will describe PhotoCity (joint work with Kathleen Tuite, Dun-Yu Hsaio, and Zoran Popovic), an project that represents a first step towards leveraging this camera by involving people in a competitive, online game where players submit photos contributing to an ever-expanding 3D model.

Keywords: Collaborative games, 3D reconstruction, structure from motion, crowdsourcing

Why points, 3D models and difficult computation?

Henrik Stewenius (Google Switzerland - Zürich, CH)

Much of the research for 3D computation is about computing accurate 3D models and pointclouds.

Most of the "models" today presented to users are not renderings of models but actual images. Most notable are Google Streetview, Bing Obliques but also Photosynth. The question I am trying to expose is: What do we compute and why?

Maps for Cognitive Automobiles

Christoph Stiller (KIT - Karlsruhe Institute of Technology, DE)

Cooperative Cognitive Automobiles shall perceive their environment from data of various sensors, use vehicle2x communications and, employ map information.

The required information needs a high level of integrity, completeness, actuality, and accuracy. The talk will outline the benefits of and requirements on maps for this application and provide some ideas for their architecture.

Keywords: Cooperative cognitive automobiles, semantic maps

Full Paper:
http://www.cognimobil.org

Detecting inconsistencies in the 3D Geometry of a scene

Aparna Taneja (ETH Zürich, CH)

A dynamic map can be visualized as an environment that changes with time. The changes in an environment can be classified as fast or slow changes. For instance, fast changes refer to a scene changing continously with time, such as people walking on a street or cars moving, and the slow changes correspond to changes that take place over a longer period of time such as a new building on a street.

We are interesting in detecting these changes in the geometry of an environment using some images that represent its current state.

We first consider the case of modeling the fast changes in a scene, and present the method proposed in our work titled "Modelling dynamic scenes recorded with freely moving cameras" [Aparna Taneja, Luca Ballan, Marc Pollefeys (ACCV'10)] where we presented a probabilistic framework to leverage the existing 3D geometry of the scene and some videos recorded with moving cameras in outdoor environments to detect geometric inconsistencies and provide a volumetric reconstruction of all the dynamic elements of the scene.

As a work in progress we also present our efforts on extending this work to be able to detect the slow changes in the geometry of a scene over time. A proposed application is to detect geometric changes in the 3D models of cities, such that the process of updating these models can be significantly optimized by limiting the update to only those areas where a change is detected.

Keywords: 3D change detection

Feature extraction from point clouds

George Vosselman (ITC - Enschede, NL)

High quality point clouds are nowadays acquired by laser scanning or dense surface reconstruction from imagery. While point clouds are suitable for visual inspection tasks, other applications require meaningful objects that can be queried. I'll present our work on the extraction of 3D building models from airborne laser scanning data as well as the extraction of street furniture from mobile laser scanning data. For the building reconstruction we've tried to find a proper balance between data and model driven approaches. Success rates are between 80 and 95%, depending on the building complexity. To get higher success rates, we feel a better modelling of knowledge about buildings (or the road environment in case of road inventory studies) is required.

Keywords: 3D reconstruction, feature extraction, modelling, point clouds

StyP-Boost: A Bilinear Boosting Algorithm for Learning Style-Parameterized Classifiers

Jonathen Warrell (Oxford Brookes University, GB)

We introduce a novel bilinear boosting algorithm, which extends the multi-class boosting framework of JointBoost to optimize a bilinear objective function. This allows style parameters to be introduced to aid classification, where style is any factor which the classes vary with systematically, modeled by a vector quantity. The algorithm allows learning to take place across different styles. We apply this Style Parameterized Boosting framework (StyP-Boost) to two object class segmentation tasks: road surface segmentation and general scene parsing. In the former the style parameters represent global surface appearance, and in the latter the probability of belonging to a scene-class. We show how our framework improves on 1) learning without style, and 2) learning independent classifiers within each style. Further, we achieve state-of-the-art results on the Corel database for scene parsing.

Keywords: Segmentation, Learning, Object Recognition

Scene Reconstruction and Understanding by Mining

Ruigang Yang (University of Kentucky, US)

We are capturing visual data of our world at an astonishing rate. How to find the structural semantics from raw pixel or point set is a important task.

Semantics are often manifested as repeated structures or patterns. Therefore we advocate to embrace the large amount of visual data and utilize data mining techniques to find repetitive patterns automatically.

I will talk about a novel unsupervised learning algorithm to find pairwise association of visual elements that have consistent geometric relationships sufficiently often. Furthermore, high-order structural semantics are extracted by mining patterns that are composed of pairwise spatially consistent association of visual elements. I will demonstrate the effectiveness of our approach with a number of data sets.

Keywords: 3D Modeling, data mining, data fusion