07171 Abstracts Collection
Visual Computing - Convergence of Computer Graphics and Computer Vision
— Dagstuhl Seminar —

Markus Gross[1], Heinrich Müller[2], Hans-Peter Seidel[3] and Harry Shum[4]

[1] ETH Zürich, CH
grossm@inf.ethz.ch
[2] Univ. Dortmund, DE
[3] MPI für Informatik, DE
[4] Microsoft Research - Beijing, CN

Abstract. From 22.04. to 27.04.2007, the Dagstuhl Seminar 07171 "Visual Computing - Convergence of Computer Graphics and Computer Vision" was held in the International Conference and Research Center (IBFI), Schloss Dagstuhl. 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. Image- and video-based modeling and rendering, perception-guided modeling and rendering, texture synthesis, scattering and reflectance measurement rendering, capturing reality (appearance, motion) from images, 3D acquisition and display, 3D reconstruction, image and model compression, computation

07171 Summary – Visual Computing – Convergence of Computer Graphics and Computer Vision

Due to the importance of visual information for humans, visual computing is at the very core of the technologies enabling the modern information society. New and emerging technologies such as multimedia, digital television, telecommunication and telepresence, or virtual reality further indicate the tremendous potential of visual interaction with computers in the years to come. Typical for the field is the coincidence of very large data sets with the demand for fast, if possible interactive, user-adapted high quality visual display of the results. Furthermore, the user should be able to interact with the environment in a natural and intuitive way. In order to address the challenges mentioned above, a new and more integrated scientific view of Visual Computing is required that unifies the previously separate "visual" disciplines of computer graphics and computer

vision. Computer graphics is traditionally concerned with generating visual interfaces of computers and applications to the user. Computer vision focuses on enabling computers to understand and interpret visual information from static images and video sequences.

Keywords: Image- and video-based modeling and rendering, perception-guided modeling and rendering, texture synthesis, scattering and reflectance measurement rendering, capturing reality (appearance, motion) from images, 3D acquisition and display, 3D reconstruction, image and model compression, computation

Joint work of: Gross, Markus; Müller, Heinrich; Seidel, Hans-Peter; Shum, Harry

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2008/1503

Extracting the essence from sets of images

Marc Alexa (TU Berlin, D)

We use a set of photographs taken from similar viewpoints as a model for a single photograph from this viewpoint. A distance measure for this model is defined by correlating the neigborhoods of pixels in similar positions. A cross analysis of the source images yields confidence values for their pixels. The confidence values together with the distances in pixels are used to steer a variable bandwidth mean shift algorithm that moves an arbitrary image towards one conforming with the model. Furthermore, distances are also used for a non-local means reconstruction of image areas that have no consistent explanation in the source images. This allows reconstructing images of scenes that are occluded in the majority of images.

Keywords: Picture and Image Generation

Graphics-based Vision

Volker Blanz (Universität Siegen, D)

In recent years, Computer Graphics has adopted more and more techniques and components from Computer Vision. One of the reasons for this is that research in Graphics has focussed a lot on real-world measurements, such as images or 3D scans, which are the domain of Computer Vision. We are now starting to see the opposite direction of transfer: For example, analysis-by-synthesis techniques rely on Computer Graphics for Vision tasks. This talk briefly summarizes an analysis- by-synthesis approach to face recognition, and then turns to a more recent, conceptually different instance of Graphics for Vision by demonstrating how synthetic front views of faces, which are generated by a Graphics system, can enhance the robustness of traditional, image-based face recognition systems.

Keywords: Face Recognition , 3D Reconstruction

Full Paper:
http://mi.informatik.uni-siegen.de/publications/cvpr05_blanz.pdf

See also: Face Recognition based on Frontal Views generated from Non-Frontal Images. Blanz, V., Grother, P. Phillips, J. and Vetter, T. IEEE Conf. on Computer Vision and Pattern Recognition CVPR 2005

Shape Deformation Based on Rigid Cells

Mario Botsch (ETH Zürich, CH)

I will present our recent work on interactive deformation of surfaces and solids. The models are embedded in a set of volumetric cells, which are coupled by a physically-inspired non-linear elastic energy. To deform the model, the cells are rigidly transformed to satisfy the user's modeling constraints while minimizing the elastic energy. The rigidity of the cells prevents degenerations even under extreme deformations, making the method numerically stable. The underlying geometric optimization is based on both local and global shape matching, and we employ hierarchical solvers and adaptive refinement strategies for improved performance.

Keywords: Shape deformation, geometric modeling

Visual recognition: How can we learn complex models?

Joachim Buhmann (ETH Zürich, CH)

Vision with its grand challenge of general scene understanding requires hierachically structured, modular representations of image content. The models also have to capture the statistical nature of images with their enormeous variability and semantic richness. The design of such models requires to resolve the tradeoff between flexibility with the resulting high model complexity on the one hand and the limited amount of data on the other hand.

Compositionality as a design principle advocates a representation scheme of image content which detects local parts like wheels for cars or eyes for faces and composes these information pieces to combinations of parts in a recursive manner. Graphical models can express both the probabilistic nature of features as well as their spatial relations. The composition machine for natural image parsing explains part relations in a generative way and allows us to synthesize images from the learned image structure. State-of-the-art categorization results both for still images and for video are achieved with a significantly more compact representation than employed by alternative approaches. (joint work with Björn Ommer)

Keywords: Object recognition, categorization, graphical models, compositionality

Full Paper:
http://ml-pub.inf.ethz.ch/publications/papers/2006/0_ommer.eccv06.pdf

On Canonical Representations of 3D Objects with Invariants

Hans Burkhardt (Universität Freiburg, D)

Computer vision and computer graphics show more and more a need for a universal platform for 3D object modelling. However, the representation of 3D objects is not canonical and objects often occur at different spatial position and in different rotational poses. Therefore the following tasks turn out to be nontrivial:

- Compare data sets
- Search in large data bases
- Find measure of similarity
- Look for partial similarities

Identical objects belong to equivalence classes under the group of Euclidean transformations. Therefore these tasks should be performed over these equivalence classes. One way to reach this goal is to compute and to use an invariant representation for all entities which e.g. then easily allows to build an object specific but pose independent index for a database.

This talk gives an overview about our recent work in 3D invariant feature representation and object retrieval. The concept of group integration is used to generate invariant signatures for the database entries [1]. We propose a method to compute such group integration features very quickly by a randomized approach and uncover the relation of such features to approaches using invariant histograms [2]. Further it is shown how to enhance the group integration approach based on the harmonic analysis of the 3D rotation group [3]. Results for 3D surface models and structural protein databases are presented.

Joint work of: Burkhardt, Hans; Reisert, Marco

Using the irreducible Representations of the 3D Rotation Group for Analysis, Synthesis and Recognition of 3D Images

Nikolaos Canterakis (Universität Freiburg, D)

The representation theory of the 3D rotation group SO(3) offers a potential for a systematic analysis of 3D images, especially when their semantic classes remain unchanged under 3D rotations.

We will describe how the construction of invariants for recognition is facilitated by the theory in this case and derive 3D Zernike invariants, from which an object can be synthesized in a canonical rotational pose.

Computational Photography

Frédo Durand (MIT - Cambridge, USA)

Computational Photography and Bilateral Image Decomposition

Fredo Durand MIT CSAIL http://people.csail.mit.edu/fredo/

The digital photography revolution has greatly facilitated the way in which we take and share pictures. However, it has mostly relied on a rigid imaging model inherited from traditional photography. Computational photography and video go one step further and exploit digital technology to enable arbitrary computation between the light array and the final image or video. Such computation can overcome limitations of the imaging hardware and enable new applications. It can also enable new imaging setups and postprocessing tools that empower users to enhance and interact with their images and videos.

This talk describes new imaging architectures as well as software techniques that leverage computation to facilitate the extraction of information and enhance images. In particular, I will describe the use of a bilateral decomposition of images into a large-scale and a detail component using an edge-preserving approach. I will describe a variety of techniques that build on such decomposition for tone mapping, relighting, style transfer and flash photography. Finally, I will describe a new data structure, the bilateral grid, which naturally enables edge-preserving image manipulations by lifting images into a higher-dimensional space.

Keywords: Computational Photography, Bilateral Filter

The Influence of Shape on the Perception of Material Reflectance

Philip Dutré (Katholieke Universiteit Leuven, B)

Visual observation is our principal source of information in determining the nature of objects, including shape, material or roughness.

The physiological and cognitive processes that resolve visual input into an estimate of the material of an object are influenced by the illumination and the shape of the object. This affects our ability to select materials by observing them on a point-lit sphere, as is common in current 3D modeling applications.

In this talk we present an exploratory psychophysical experiment to study various influences on material discrimination in a realistic setting. The resulting data set is analyzed using a wide range of statistical techniques. Analysis of

variance is used to estimate the magnitude of the influence of geometry, and fitted psychometric functions produce significantly diverse material discrimination thresholds across different shapes and materials.

Suggested improvements to traditional material pickers include direct visualization on the target object, environment illumination, and the use of discrimination thresholds as a step size for parameter adjustments.

Keywords: Visual perception, psychophysics, shading, material

Model Extraction for Computer Graphics

Eugene Fiume (University of Toronto, CA)

Many models in computer graphics are intended in some way to mimic nature.

As such, it is sensible to try to extract salient parameters of realistic models from real-world observations. For many directly observable phenomena such as reflectance, texture, morphology and kinematics, this is relatively easy to accomplish, although retargetting real-world data is not always easy. Indirectly observed phenomena such as wind, friction, brittleness, plasticity, and dynamics, however, pose a particular challenge. As a demonstration, we discuss how to extract a physical wind field from a single video sequence of swaying trees. A physically based solution to this problem requires inverting the usual equations of motion (harmonic oscillation) to compute the forces that account for an observed displacement, and a second inversion to compute the wind field deforming a swaying tree. While successful in a purist sense, the complexity of this solution suggests that hybrid approaches should also be explored. We discuss some recent work in which the motion of trees extracted from a single, two-dimensional video sequence is retargetted to three-dimensional synthetic trees without any appeal to physics.

These two approaches serve as useful bookends for the discussion of the tools that need to developed for general model extraction problems.

Keywords: Model extraction, computer animation, physically-based models, computer vision, computer graphics, harmonic oscillation, inverse dynamics

Rendering for Tracking - A Perspective from Computer Vision

Jürgen Gall (MPI für Informatik - Saarbrücken, D)

While methods from computer vision are commonly applied in computer graphics, the potential of computer graphics is mostly untapped for computer vision tasks. For problems like visual motion capture, however, the appearance of the observed object is an essential cue that can be only efficiently integrated by techniques known from computer graphics. In this talk, the use of rendering

techniques for tracking is discussed by means of incorporating the appearance of a 3D model. Hence, rendering is not regarded as generating realistic-looking models but as an efficient representation of relevant texture information of the model's surface. Several examples with articulated models of humans with up to 30 degrees of freedom are shown to demonstrate the progress that can be achieved by incorporating rendering techniques into a vision-based framework.

Multi-View Stereo Revisited

Michael Goesele (University of Washington, USA)

I will present an extremely simple yet robust multi-view stereo algorithm and analyze its properties. The algorithm first computes individual depth maps using a window-based voting approach that returns only good matches. The depth maps are then merged into a single mesh using a straightforward volumetric approach. I will show results for several datasets, showing accuracy comparable to the best of the current state of the art techniques and rivaling more complex algorithms. I will also give an overview over ongoing work on multi-view stereo reconstruction for more complex and general datasets.

Improving Dynamic 3D Scanning with Image Based Motion Compensation

Stefan Gumhold (TU Dresden, D)

Structured light approaches based on stripe patterns allow for 3D scanning with high resolution, simple implementation and robust results in the static case. Due to the necessity of several image captures for one 3D scan, the dynamic case is significantly harder. Movement of dynamic object surfaces induces movement of the stripe patterns. As a result the correspondences between patterns in different captured images are destroyed if the object moves too fast. We propose a new image based technique to estimate and compensate pattern motions resulting in a robust 3D scanning approach of dynamic objects.

Keywords: Dynamic 3d scanning, motion compensation

Joint work of: König, Sören; Gumhold, Stefan

Image-Based Motion Compensation for Structured Light Scanning of Dynamic Surfaces

Stefan Gumhold (TU Dresden, D)

Structured light scanning systems based on temporal pattern codification produce dense and robust results on static scenes but behave very poorly when applied to dynamic scenes in which objects are allowed to move or to deform during the acquisition process.

The main reason for this lies in the wrong combination of encoded correspondence information because the same point in the projector pattern sequence can map to different points within the camera images due to depth changes over time.

We present a novel approach suitable for measuring and compensating such kind of pattern motion.

The described technique can be combined with existing active range scanning systems designed for static surface reconstruction making them applicable for the dynamic case. We demonstrate the benefits of our method by integrating it into a gray code based structured light scanner, which runs at thirty 3d scans per second.

Keywords: 3d scanning, motion compensation, optical flow, structured light, dynamic surfaces

Joint work of: Gumhold, Stefan; König, Sören

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2008/1502

Visual Computing with the Human in the Loop

Marc Hanheide (Universität Bielefeld, D)

Visual Computing is widely seen as the convergence of computer graphics and computer vision. In interactive systems, however, a human user joins this symbiosis; imposing certain challenges and chances for the system's operation. A tight coupling between system and human user allows for efficient and effective collaboration. In my talk, I present work conducted in the context of the EU project VAMPIRE to develop a wearable assistance system that embeds the user in the visual processing loop by means of head-mounted cameras and augmented reality. Utilizing this shared perception and mutual control system and user can supervise and correct each other in visually demanding tasks. The presented assistance system is enabled to perceive and understand objects and actions in the field of view of the user and provide assistance in object manipulation tasks; discussed at the example of a cocktail preparation scenario.

Keywords: Cognitive vision, augmented reality, assistance system, action recognition, computer vision

Full Paper:
http://bieson.ub.uni-bielefeld.de/volltexte/2007/1032/

See also: @phdthesisHanheide2006-ACE, author = Marc Hanheide, title = A Cognitive Ego-Vision System for Interactive Assistance, school = Technische Fakultät – Universität Bielefeld, year = 2006, month = dec, keywords = Assistance systems, Cognitive vision, Augmented reality, owner = mhanheid, pdf = Hanheide2006-ACE.pdf, timestamp = 2007.02.19, volltext = http://bieson.ub.uni-bielefeld.de/volltexte/2007/1032/

Shape Design is Rule Design

Sven Havemann (TU Graz, A)

Most man-made shapes are actually representatives of whole shape families, generated by generalizable constructions. This is reason why parametric modeling has become the mainstream approach today for the industrial design of three-dimensional shapes. There is, however, a multitude of different, and mostly incompatible, approaches. I will present our efforts to design an architecture for a general procedural exchange format. But I will also highlight a number of unsolved problems. One of the most promising, but also most difficult problems is Generative Surface Reconstruction, which is to find a suitable parametrization of a given shape. This would permit, e.g., to re-sythesize the construction plans for a digitized industrial 3D object.

Keywords: 3D modeling, shape design, parametric design, generative modelling

Towards a High Dynaic Range Imaging Pipeline

Wolfgang Heidrich (University of British Columbia - Vancouver, CA)

The power of the human visual system to process wide ranges of intensities far exceeds the abilities of current imaging systems. Both cameras and displays are currently limited to a dynamic range (contrast) of between 300:1 to 1,000:1, while the human visual system can process a simultaneous dynamic range of 50,000:1 or more, and can adapt to a much larger range.

High-dynamic-range (HDR) imaging refers to the capture, processing, storage, and display of images with significantly improved contrast and brightness compared to the conventional imaging pipeline. This new HDR imaging pipeline is designed to match the power of the human visual system. HDR displays significantly improve the sense of realism and immersion when showing both real and synthetic HDR imagery. Likewise, HDR cameras are able to take images without saturation under difficult lighting situations. The additional information captured in both extremely bright and extremely dark regions is useful as an input for HDR displays, but also for machine vision applications.

In this talk, I will summarize the results of a multi-disciplinary collaboration between Physics and Computer Science researchers at the University of British Columbia, as well as the UBC spinoff company BrightSide technologies, which was recently acquired by Dolby Labs.

Physics-Based Human Motion Models for Animation and Tracking

Aaron Hertzmann (University of Toronto, CA)

I describe physics-based models of human motion, with applications to computer animation and 3D person tracking.

I begin by surveying relevant principles of motion from the biomechanics literature. I then describe a representation of motion that incorporates passive dynamical elements, relative muscle preferences, and optimality principles. I then describe Nonlinear Inverse Optimization, a novel learning algorithm that can be used to learn these physical models of motion from data. Once learned, these models can then be synthesize new animation. I then describe a tracking algorithm for on-line estimation of 3D human motion from video sequences that makes use of a physics-based prior model of motion.

Full Paper:
http://www.dgp.toronto.edu/~hertzman/

See also: Liu, C. K., Hertzmann, A. and Popović;, Z. Learning Physics-based Motion Style with Nonlinear Inverse Optimization SIGGRAPH 2005.; M Brubaker, D Fleet, A Hertzmann. Physics-Based Person Tracking Using Simplified Lower-Body Dynamics

Reconstruction and Rendering of Time-Varying Inhomogenous Refractive Index Fields

Ivo Ihrke (MPI für Informatik - Saarbrücken, D)

We present a technique for 2D imaging and 3D tomographic reconstruction of time-varying, inhomogeneous refractive index fields.

Our method can be used to perform three-dimensional reconstruction of phenomena such as gas plumes or liquid mixing. We can also use the 2D imaging results of such time-varying phenomena to render environment mattes and caustics.

To achieve these results, we improve a recent fluid imaging technique called Background Oriented Schlieren imaging, and develop a novel theory for tomographic reconstructions from Schlieren images based on first principles of optics. We demonstrate our approach with three different measurement setups, and discuss example applications such as measuring the heat and density distribution in gas flows.

Additionally, we present an interactive rendering technique for the acquired phenomena.

Render-based Vision - The analysis by synthesis approach for model-based computer vision

Reinhard Koch (Universität Kiel, D)

The Analysis by Synthesis approach in model-based computer vision has long since relied on computer graphics rendering to facilitate parameter estimation.

Examples include model-based tracking and shape estimation amongst others. With the advent of powerful graphics engines this approach can even be pushed further. Complex 3D scenes can be rendered and compared with real images online, giving rise to realtime camera tracking and model fitting even with general camera models like fisheye lenses. In this talk I will discuss the use graphics to facilitate computer vision by exploiting the paradigm of render-based vision.

Keywords: Graphics and vision, Model-based computer vision, analysis by synthesis approach

Superresolution in Reflectance Fields

Hendrik Lensch (MPI für Informatik - Saarbrücken, D)

Captured reflectance fields provide in most cases a rather coarse sampling of the incident light directions. As a result, sharp illumination features, such as highlights or shadow boundaries, are poorly reconstructed during relighting; highlights are disconnected, and shadows show banding artefacts. In this paper, we propose a novel interpolation technique for 4D reflectance fields that is capable of reconstructing plausible images even for non-observed light directions. Given a sparsely sampled reflectance field, we can effectively synthesize images as they would have been obtained from denser sampling.

The processing pipeline consists of three steps: segmentation of regions where apparent motion cannot be obtained by blending, appropriate flow algorithms for highlights and shadows, plus a final reconstruction technique that uses image-based priors to faithfully correct errors that might be introduced by the segmentation or flow step. The algorithm reliably reproduces scenes containing specular highlights, interreflections, shadows or caustics.

Solid Texture Synthesis from 2D Exemplars

Dani Lischinski (The Hebrew University of Jerusalem, IL)

We present a novel method for synthesizing solid textures from 2D texture exemplars. First, we extend 2D texture optimization techniques to synthesize 3D texture solids. Next, the non-parametric texture optimization approach is integrated with histogram matching, which forces the global statistics of the synthesized solid to match those of the exemplar. This improves the convergence of the synthesis process and enables using smaller neighborhoods. In addition to producing compelling texture mapped surfaces, our method also effectively models the material in the interior of solid objects. We also demonstrate that our method is well-suited for synthesizing textures with a large number of channels per texel.

Joint work of: Kopf, Johannes; Fu, Chi-Wing; Cohen-Or, Daniel; Deussen, Oliver; Lischinski, Dani; Wong, Tien-Tsin

Full Paper:
http://johanneskopf.de/publications/solid/index.html

The Exact, the Consistent, and the Credible - degrees of freedom in Visual Computing

Marcus Magnor (TU Braunschweig, D)

To deviate from the rigor of physical laws whenever deemed advantageous is a long-standing tradition in computer graphics.

Paradoxially, yet fortunately, even "un-natural" images may be perceived as being truly authentic renditions of the real world. This latitude in human visual perception greatly expands the options of image-based modeling and rendering techniques. In my talk, I'll discuss different examples of how visual computing benefits from demanding any of the following: unique exactness, physical consistency, or perceptional credibility.

Keywords: Perception, image-based, rendering, modeling

Contrast restoration by adaptive countershading

Karol Myszkowski (MPI für Informatik - Saarbrücken, D)

We address the problem of communicating contrasts in images degraded with respect to their original due to processing with Computer Graphics algorithms. Such degradation can happen during tone mapping of High Dynamic Range images, or while rendering scenes with low contrast shaders or poor lighting. Inspired by a family of known perceptual illusions: Craik-O'Brien-Cornsweet, we enhance contrasts by modulating brightness at edges to create countershading profiles.

We generalize unsharp masking by coupling it with a multi-resolution local contrast metric to automatically create countershading profiles from sub-band components which are individually adjusted to each corrected feature to best enhance contrast with respect to the reference. Additionally, we employ a visual detection model to assure that our enhancements are not perceived as objectionable halo artifacts. The overall appearance of images remains mostly unchanged and the enhancement is achieved within the available dynamic range.

We use our method to post-correct tone mapped images and improve images using depth information.

Joint work of: Krawczyk, Grzegorz; Myszkowski, Karol; Seidel, Hans-Peter

Everything You Always Wanted to Know About Sex (perception) * But Were Afraid to Ask

Carol O'Sullivan (Trinity College - Dublin, IRL)

And now that I've got your attention....

In order to create realistic images and animations of Virtual Humans, or to accurately recognise real humans from photos or video, an understanding of the factors affecting perception of sex (i.e., male, female) is required. I will discuss some recent research results on the role of shape and motion in the perception of sex, gender (i.e., masculinity and femininity) and attractiveness - and I will explain why we care about this for simulating large crowds of virtual humans....

(and apologies to Woody Allen for abusing his movie title)

Visual Computing in Sensorimotor Biology

Dinesh Pai (University of British Columbia - Vancouver, CA)

Human character animation currently relies on ad hoc and labor intensive techniques. Until recently the physical environment was also animated in this way, but has now been dramatically transformed, with ad hoc methods replaced by increasingly sophisticated and powerful *simulation* methods. I believe human character animation is also poised to transform into human simulation in a similar way. Animations could be specified at a much higher level, with the capability to automatically reproduce important features such as the visible movement of muscles, tissue deformation on contact, and dynamic constraints. These techniques will also contribute to a deep and constructive understanding of the principles of sensorimotor biology. In this keynote lecture, I will first review the past work of my group in *interactive simulation* and *reality-based modeling*, and then describe our current work in constructing detailed models of human sensorimotor biology.

Keywords: Sensorimotor, biology, musculoskeletal, neuroscience, reality-based, interactive simulation

Symmetry Detection and Symmetrization

Mark Pauly (ETH Zürich, CH)

"Symmetry is a complexity reducing concept [...]; seek it everywhere" Alan J. Perlis. -

I will discuss our work on finding and enhancing partial and approximate symmetries in 2D and 3D geometry. Local symmetry evidence based on pair-matchings of surface samples is accumulated in an appropriate transformation space.

Clustering methods are applied to extract relevant symmetry transforms, including translation, rotation, and uniform scaling. An optimization approach that couples the spatial domain with the symmetry transformation space can be formulated to symmetrize geometric shapes.

Keywords: Symmetry, shape optimization

Visual 3D Modeling of the Real World

Marc Pollefeys (University of North Carolina- Chapel Hill, USA)

Images and videos form a rich source of information about the visual world. The extraction of 3D information from images is an important research problem in computer vision and graphics. The ubiquitous presence of cameras and the tremendous advances of processing and communication technologies yields important opportunities and challenges in those areas.

My work has focused on developing flexible techniques for recovering 3D shape, motion and appearance from images. A first example of this is an approach to recover photo-realistic 3D models of static objects or scenes from videos recorded with a hand-held camera or on a moving vehicle. A key aspect of our approach is the ability to also recover the geometric and photometric calibration of the camera from the image data so that our techniques can also work with uncalibrated consumer cameras or archive photographs. Towards the end of my talk, I will also briefly discuss approaches to capture dynamic scenes, both from single and multiple cameras. Applications ranging from archaeology and 3D urban modeling, to special effects and 3D tele-medecine will be used to illustrate our work.

Keywords: Computer Vision, 3D Modeling, Structure from Motion, Multi-View Stereo

Full Paper:
http://video.google.com/videoplay?docid=-1315387152400313941

Model Driven Motion Capture as Example Towards Graphics for Vision

Bodo Rosenhahn (MPI für Informatik - Saarbrücken, D)

The presentation deals with aspects on model based computer vision. The focus is on integration of prior knowledge e.g. modeling techniques known from computer graphics in tasks for computer vision.

Several examples from the field of markerless human motion capture are given to show the improvements, e.g. in the fields of image segmentation, pose tracking, tracking in highly noisy environments, bridging frame drops or cloth modeling and tracking clothed people.

High-order Markov Random Fields for Low-Level Vision

Stefan Roth (Brown Univ. - Providence, USA)

A large number of problems in low-level vision can be approached using probabilistic inference. One key component of these approaches is modeling the prior distribution of the dense scene representation that is to be recovered. In image restoration applications, such as in image denoising, this amounts to modeling the prior probability of having a particular image among all possible images in the world. Markov random fields (MRFs) have found widespread use for modeling such prior knowledge, but in the past have mostly relied on simple random field structures that only model interactions between neighboring pixels. I will argue that these pairwise MRFs are not powerful enough to capture the rich statistics of natural images and other dense scene representations. To overcome this, I will introduce a new high-order Markov random field model, termed Fields of Experts (FoE), that better captures the structure of natural images by modeling interactions among larger neighborhoods of pixels. The parameters of this model are learned from a database of natural images using contrastive divergence. One of the key advantages of such a generic, but powerful prior model is that it can be applied to a wide variety of different applications. I will demonstrate the capabilities of the FoE model with image denoising and inpainting applications.

Stochastic Optimization of Multiple Texture Registration

Dietmar Saupe (Universität Konstanz, D)

We consider the problem of simultaneously registering sev- eral images to a 3D model. We propose a global approach based on mutual information that extends previous methods to incorporate the color, and does not require segmentation or feature extraction. We give a stochastic model for joint optimization of multiple image-to-model align- ment and we propose a heuristic to solve it. Experiments with synthetic models showed that our algorithm is robust to illumination and to vary- ing surface characteristics. Experiments with real data showed that we can achieve very good precision even for an ob ject with highly specular surface, in moderate lighting environment.

What do you need to scan your city? The Tübingen scanning project

Andreas Schilling (Universität Tübingen, D)

The Wägele, Analysis-by-Synthesis, Generative Models, Bayes, how they relate to each other and other questions in the context of scanning and modelling from acquired data.

Towards Compact Tele-Immersion

Oliver Staadt (Univ. of California - Davis, USA)

Tele-Immersion technology enables users at geographically distributed sites to collaborate in real time in a shared, simulated environment as if they were in the same physical space. This new approach for human-computer interaction is the synthesis of networking and media technologies. Prior projects, such as the blue-c and the National Tele-Immersion Initiative developed prototype systems that addressed some of the challenges, but also put forth unresolved issues and the need for further research. In this talk, I will present our vision of a novel compact tele-immersion system that will eventually support bidirectional communication and interaction between users located at more that two geographic sites.

Receiver- and Channel-adaptive Compression for Remote Browsing of Image-Based Scene Representations

Eckehard Steinbach (TU München, D)

Remote navigation in compressed image-based scene representations requires random access to arbitrary parts of the reference image data to recompose virtual views. The degree of inter-frame dependencies exploited during compression has an impact on the effort needed to access reference images and delimits the rate distortion (RD) trade-off that can be achieved.

This work considers conventional RD optimization but additionally takes a given receiver hardware and a maximum available transmission bitrate into account. This leads to an extension of the traditional rate-distortion optimization to a trade-off between the four parameters rate (server side file size), distortion, transmission data rate, and decoding complexity. This RDTC optimization framework allows us to adapt to channel properties and client resources and can significantly improve the user satisfaction in a remote navigation scenario.

Keywords: Image-based rendering, interactive streaming, RDTC optimization

Joint work of: Steinbach, Eckehard; Bauermann, Ingo

3D Video Billboard Clouds

Michael Waschbüsch (ETH Zürich, CH)

3D video billboard clouds reconstruct and represent a dynamic three-dimensional scene using displacement-mapped billboards. They consist of geometric proxy planes augmented with detailed displacement maps and combine the generality of geometry-based 3D video with the regularization properties of image-based 3D video.

3D video billboards are an image-based representation placed in the disparity space of the acquisition cameras and thus provide a regular sampling of the scene with a uniform error model. We propose a general geometry filtering framework which generates time-coherent models and removes reconstruction and quantization noise as well as calibration errors. This replaces the complex and time-consuming sub-pixel matching process in stereo reconstruction with a bilateral filter. Rendering is performed using a GPU-accelerated algorithm which generates consistent view-dependent geometry and textures for each individual frame. In addition, we present a semi-automatic approach for modeling dynamic three-dimensional scenes with a set of multiple 3D video billboards clouds.

Joint work of: Waschbüsch, Michael; Würmlin, Stephan; Gross, Markus

Beauty with Variational Methods: An Optic Flow Approach to Hairstyle Simulation

Joachim Weickert (Universität des Saarlandes, D)

In this talk we show how well-adapted variational ideas can solve the problem of hairstyle simulation in a fully automatic way: A customer in a hairdresser's shop selects a new hairstyle from a database, and this hairstyle is automatically registered to a digital image of the customer's face. Interestingly already a carefully modified optic flow method of Horn and Schunck turns out to be ideal for this application.

These modifications include an extension to colour sequences, an incorporation of warping ideas in order to allow large deformation rates, and the inclusion of shape information that is characteristic for human faces. Employing classical numerical ideas such as finite differences and SOR iterations offers sufficient performance for real-life applications. In a number of experiments we demonstrate that our variational approach is capable of solving the hairstyle simulation problem with high quality in a fully practical setting.

Joint work of: Weickert, Joachim; Demetz, Oliver; Bruhn, Andres; Welk, Martin

Reassembling the Thera Frescoes

Tim Weyrich (Princeton University, USA)

The archaeological site of Akrotiri on the island of Thera (modern-day Santorini) has proven a treasure trove of information about prehistoric Aegean civilization and culture.

Among the most valued excavation finds are wall paintings (frescoes), which have been preserved in the volcanic ash since the sixteenth century BCE. However, the frescoes are typically recovered in fragments of a few centimeters to

a few tens of centimeters in length, and reconstructing complete wall sections from the fragments occupies a major portion of the effort at Akrotiri. In this talk I will present an ongoing project, in which we are building a system that will assist archaeologists by digitizing excavated fragments and automatically proposing matches on the basis of color, 3-D shape, and other cues. The goal is a more rapid and more complete reconstruction of the wall paintings, with less handling of the physical fragments. I will talk about the current state of the frescoes and how they are reconstructed and conserved, as well as our ideas for building computer-based tools to help the process.

A Finite Element Method on Convex Polyhedra

Martin Wicke (ETH Zürich, CH)

Finite element methods are a tremendously important tool in computer animation. Tetrahedral discretizations are the standard in this field. After topological changes in the simulation domain, maintaining a good tetrahedral discretization is challenging.

Removing the restriction of tetrahedral decompositions and allowing arbitrary convex polyhedral elements yields a more flexible simulation method. Our finite element approach draws upon recently introduced 3D mean value coordinates to define smooth interpolants within the elements. The mathematical properties of our basis functions guarantee convergence. Our method is a natural extension to linear interpolants on tetrahedra: for tetrahedral elements, the methods are identical. For fast and robust computations, we use an elasticity model based on Cauchy strain and stiffness warping.

This more flexible discretization is particularly useful for simulations that involve topological changes, such as cutting or fracture. Since splitting convex elements along a plane produces convex elements, remeshing or subdivision schemes used in simulations based on tetrahedra are not necessary, leading to less elements after such operations. We propose various operators for cutting the polyhedral discretization. Our method can handle arbitrary cut trajectories, and there is no limit on how often elements can be split.

There are open questions regarding the accuracy of the method compared to traditional FEM methods, feasibility for nonlinear elasticity, and the optimal formulation of the basis functions.

Keywords: Physically-based animation, elasticity simulation, finite element method, mean value coordinates