Research Statement Oliver Grau

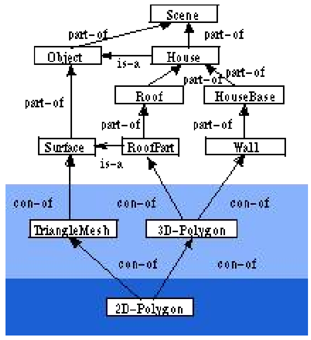
1 Introduction

My research interests have originally evolved from work on pattern recognition and scene understanding. Early projects in my scientific career were investigating pattern recognition techniques for industrial applications. This work included work on car license plate recognition and investigations on the use of neural networks for medical applications. From my PhD work onwards my effort was dedicated to visual computing; that is the combination of vision and graphics: My PhD work was looking into using knowledge-based scene analysis to generate 3D models of buildings for use in graphics applications, e.g. for visualisation of architectural structures.
In an academic post-doc assignment I was leading an interdisciplinary team to look into the use of computer vision algorithms for visual effects production. In this work we combined structure-from-motion and camera tracking to enable new abilities to combine virtual and real scene content. I kept working on this general topic through out my career as it is addressing like no other application a number of central challenges of visual computing: Determining the mathematical parameter of cameras, by means of camera calibration and camera tracking and then establishing optical interaction of computer generated imagery (CGI) with the real world using three-dimensional models created by reconstruction. My recent work also includes the estimation and harmonization of real scene lighting with CGI.
In the following years I was able to expand my experience on visual computing in the innovative labs of BBC R&D in the UK. This gave me deep insides into the media industry and how visual computing is applied to real-world tasks. My research contributions included multi-camera performance capture, image-based rendering techniques, innovative human computer interfaces (HCI) and new user experiences, either through 3D displays or interactive content.
Recently my work is more broadly looking into visual computing as the Intel director of the Intel-Visual Computing Institute, with a keen interest in above topics and additionally emerging trends, like deep learning and perceptional rendering and computational displays.

2 Summary of previous and ongoing work

The major milestones in visual media and communications alike, have been the introduction of electronic cameras and transmission in the 20th century, which allowed transmission of scenes between different locations and groups of people (in the case of broadcast). The next step in evolution was digitization and with that the emerging visual computing. Visual computing relies on one hand on the extraction of some semantic meaningful information from imaging sensors, the processing or augmentation of this information and the visualization of the resulting data. The following text summarizes my scientific contributions to these aspects.

Semantic scene interpretation and 3D-modelling from images



The central idea of this work was to show how semantic scene models can be automatically used to improve 3D reconstruction of (static) structures like buildings from stereoscopic images [Grau95,Grau97]. In order to achieve that a frame-like knowledge representation was defined and the scene interpretation system 'AIDA' was developed (in collaboration with the aerial scene interpretation team of Hanover University). The geometrical constraints (like symmetries, planarity, etc.) were then used to generate 3D surface models from building from images. The work was awarded the 'DAGM price 1995' of the German Pattern association [Grau95] and influenced work on scene interpretation and 3D reconstruction. Advances in computing power led in recent years to a renewed interest in neural networks and progress in deep learning methods opened also the application of these techniques to depth data. The work in [Sharma16] takles the problem of incomplete or noisy data in RGB-Z cameras by applying convolutional neural networks (CNNs) trained with deep learning methods to improve 3D reconstruction.

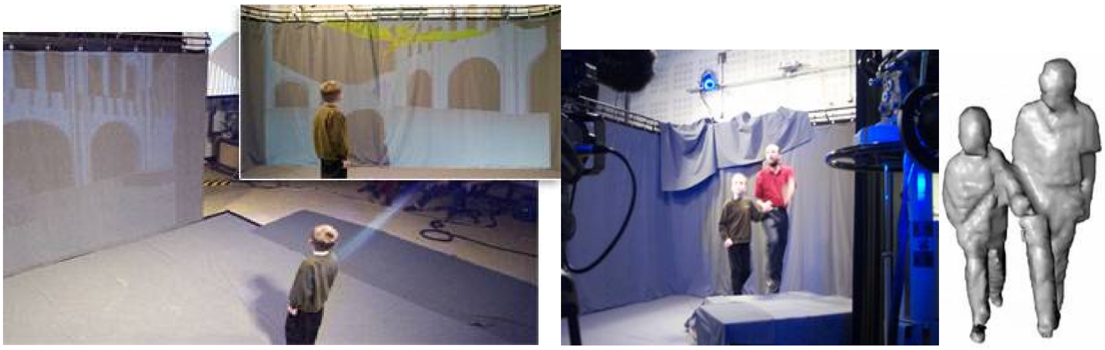
Multi-view 4D reconstruction of human action for media applications

Although 3D reconstruction of static objects is of interest for film, TV and computer games, e.g. as CGI background models, media deals with action of real actors. Visual computing in media has a number of applications, of which the production of visual effects is the most prominent. A milestone was the initiation and my work on the EU-funded ORIGAMI project (2001-2004). Subject was the development of tools for 3D reconstruction of dynamic scenes (hence 4D reconstruction) and their use in special effects and TV-productions. My scientific contributions here were the application of methods formally used in reconstruction of static objects for real-time capable reconstruction of human actors in the studio [Grau04].

The application of 4D reconstructions and free-viewpoint rendering techniques to sport events, like football is a valuable tool to sport commentary [Grau07]. The move from the defined conditions of a studio into an outdoor scenario requires more robust algorithms and methods and presented a major step. My scientific contributions here were in the development of robust segmentation techniques [Grau08], calibration error-tolerant 3D reconstruction and real-time capable distributed processing [Easter10] (see [Grau11] for an overview). Aspects of early work on the sport visualisation went into development of the 'Piero' system for the 3D visualisation of sport scenes that won the IBC Innovation Award 2006 and the Queen's Award in 2011.

Visualisation for HCI and new user experiences

For improved usability of our studio performance capture we developed a new innovative on-set visualisation system that enabled intuitive interaction with virtual content. The system combines virtual reality techniques to give actors a viewdependent visualisation of the virtual scene components, while still preserving the ability of robust chroma-keying [Grau04]. The system solved in particular the hard problem of getting the eye-gaze between the actor and virtual objects correct (eye-line problem).



The 4D capture pipeline developed for free-viewpoint applications can also be used to generate a stereoscopic user experience. In [Grau09] we have demonstrated as part of the EU-funded 3D4YOU project a production pipeline from a set of normal operating broadcast cameras to an auto-stereoscopic end-user display. In [Budd12] we presented the first streaming of free-viewpoint video into a web-browser. An approach that was extended in the Re@ct project [Grau12] to generate interactive content in a games engine.

2 Future Research

My recent research interests are on computational displays and rendering techniques for virtual and augmented reality. Although progress in graphics hardware and gaming technology promises to generate any imaginable quality of 3D content in real-time there are new constraints and requirements arising: For emerging wearable devices, like head-mounted displays (e.g. Oculus Rift) it is desired that these are running mobile, i.e. wireless. That puts serious constraints on the available graphics performance. An option would be wireless transmission of graphics from a server [Pohl15], but at the same time also the display resolution grows fast, both in spatial (pixel resolution) and sampling rate. One avenue I am particular interested to solve these problems is perceptional rendering in computational display technologies. The most prominent example of these techniques is foveated rendering that renders only image areas in highest resolution where it can be perceived by the human visual system (e.g. [Guenter12]). This area is a very promising area with many more psycho-visual effects to be exploited, like colour, peripheral frequency flicker and stereoscopic effects (patent and publication pending).

Other future work strands I am interested in beyond media are Industrial applications of visual computing. In particular concepts for the automobile of the future, that includes concepts for Advanced Driving Assistance Systems (ADAS) and in-vehicle infotainment (IVI). The challenge here is to combine a multitude of available visual sensors on one side into a visual representation of the driver and passengers that is helpful and not distracting.
Another import application area I see in tools for education and life-long training. Here the emerging VR+AR could help society to learn (e.g. in STEM education) and to keep the work-force trained in a fast changing economy. Both aspects need tools to

generate the training content, need the ability to connect to the real world and new visualisation and HCI tools.

References

[Grau97] O. Grau, ’A Scene Analysis System for the Generation of 3-D Models’, Proc. International Conference on Recent Advances in 3-D Digital Imaging and Modeling. May 12-15, 1997, Ottawa, Ontario Canada. pp 221- 228. IEEE Computer Society Press. ISBN 0-8186-7943-3.

[Grau95] O. Grau, ’Ein Szeneninterpretationssystem zur Modellierung dreidimensionaler Körper’, 17. DAGM-Symposium Mustererkennung 1995, 13.-15. Sep. 1995, Bielefeld, Germany. **Best Paper Award**.

[Sharma16] Abhishek Sharma, Oliver Grau, Mario Fritz, 'VConv-DAE: Deep Volumetric Shape Learning Without Object Labels'. (Technical Report), arXiv:1604.03755 [cs.CV], 2016.

[Grau11] O. Grau, "Multi-view 4D reconstruction of human action for entertainment applications", book chapter in ‘Visual Analysis of Humans: Looking at People’, Moeslund, Th.B.; Hilton, A.; Krüger, V.; Sigal, L. (Eds.), Springer 2011.

[Grau04] O. Grau, T. Pullen, G. Thomas, „A combined studio production system for 3D capturing of live action and immersive actor feedback“, IEEE Tr. on Systems and Circuits for Video Technology, March 2004.

[Grau07] O. Grau, G. Thomas, A. Hilton, J. Kilner, J. Starck, ’A robust free-viewpoint system for sport visualisation’, proc. of IEEE 3DTV conference, May 2007.

[Grau08] Grau, O. Easterbrook, J., 'Effects of camera aperture correction on keying of broadcast video', Proc. of IET 5th European Conference on Visual Media Production (CVMP 2008), London, UK, 26-27 Nov. 2008 , ISBN: 978-0-86341-973-7.

[Easter10] Jim Easterbrook, Oliver Grau, Peter Schubel, 'A system for distributed multi-camera capture and processing', Proc. Of CVMP 2010, London, UK.

[Grau09] O. Grau, V. Vinayagamoorthy, 'Stereoscopic 3D sports content without stereo rigs', Proc. of IBC 2009, Amsterdam, NL.

[Budd12] Chris Budd, Oliver Grau, Peter Schübel, “Web delivery of free-viewpoint video of sport events”, Proc. of 2012 NEM Summit, October 16-18 2012, ISBN 978-3-00-039710-3, Istanbul, Turkey.

[Grau12] Oliver Grau, Edmond Boyer, Peng Huang, David Knossow, Emilio Maggio, David Schneider, „RE@CT - Immersive production and delivery of interactive 3D content“, Proc. of 2012 NEM Summit, October 16-18 2012, ISBN 978-3-00-039710-3, Istanbul, Turkey.

[Pohl15] Daniel Pohl, Bartosz Taudul, Richard Membarth, Stefan Nickels, Oliver Grau, „Advanced In-home Streaming to Mobile Devices and Wearables“, In: International Journal of Computer Science & Applications (IJCSA), Vol. o.A. Pages 1-17, Technomathematics Research Foundation, 3/2015.

[Guenter12] Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, and John Snyder, Foveated 3D Graphics, ACM SIGGRAPH Asia, 2012.