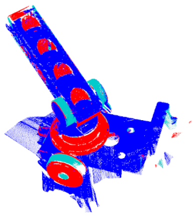
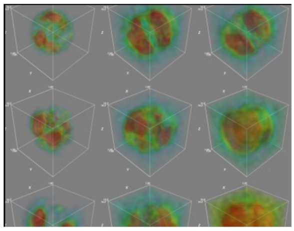
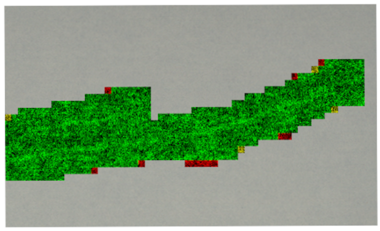
HTWG Hochschule Konstanz
Technik, Wirtschaft und Gestaltung

IOS
INSTITUT FÜR OPTISCHE SYSTEME

Institute for optical systems

annual report 2017

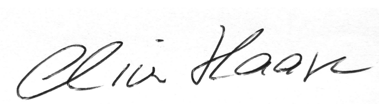
Foreword

Hochschule Konstanz University of Applied Sciences (HTWG) has been one of the first Universities of Applied Sciences that has clearly committed itself to applied research. Today, HTWG has positioned itself in the growing group of research active Universities of Applied Sciences, and is continuously striving for even greater excellence, visibility, and leadership.

The research institutes play an essential role in the overall research activities of HTWG. The Institute for Optical Systems (IOS) has been founded as one of three focus institutes in 2008, and has lived up to its mission and HTWG's expectation ever since.

Measuring the success of research is generally a difficult task. It becomes easy, however, for individuals and institutes alike that contribute to the scientific community in many different ways. The IOS clearly is such an institute. Its members have published their work in numerous high quality publications, they have acquired substantial amounts of external funding, and they have successfully led several doctoral candidates to their advanced degree, to name only the most obvious contributions. The institute's success has only been possible because of the enthusiasm of its members, and its tight integration with HTWG. Students have contributed to the research work on various levels, and the results have been fed into the members' teaching. In this sense, the IOS has demonstrated the importance of excellent applied research also as a means for high quality teaching, in particular on the graduate level.

I am proud to write these few lines of thanks and congratulations for the IOS and cordially wish the institute, but first and foremost its members, ongoing enthusiasm, thirst for knowledge, and success for the future.



Prof. Dr. Oliver Haase
Vice-President for Research, University of Applied Sciences Konstanz

Preface

The present report gives an overview over the research and development activities of the Institute for Optical Systems (IOS) Konstanz in the year 2016. Our main fields of interest are cognitive systems, geometric modelling, image processing, optical metrology and light engineering, thus representing the fundamental disciplines of current optical technology. The focus on optical systems as a whole allows us to offer competent partnership to the local industry in all relevant aspects. The IOS was founded in April 2008 by four professors from three different faculties of the University of Applied Sciences at Konstanz and is led by Prof. Dr. Umlauf (director) and Prof. Dr. Franz (associate director).

The following project descriptions present ongoing activities mainly on a status report level. Most of the reports are written by students working on their diploma, bachelor, master or Ph.D. theses. They reflect the largely varying levels of content, practice and insight that are characteristic for an institution involved in academic education. Due to its interdisciplinary nature, research at the IOS combines approaches from optics, computer graphics, image and signal processing, metrology, light engineering and sensor technology typically resulting in optical systems on a prototype level, either in pre-phase research or in cooperation with local industry.

On the occasion of our annual report, we would like to thank all of our students and co-workers for their enthusiasm and dedication which makes our institute a great place to be. Special thanks go to our institute officers, Pascal Laube, Martin Schall and Michael Grunwald for smoothly managing our day-to-day activities. We are also indebted to the administration and staff of the HTWG Konstanz for their help, especially president Dr. Carsten Manz and Prof. Dr. Oliver Haase, for their support and for continuing the start-up funding, and the faculties of Mechanical Engineering, Electrical and Information Engineering, and Computer Science with the deans Prof. Dr.-Ing. Klaus Schreiner, Prof. Dr. Thomas Birkhölzer and Prof. Dr. Jürgen Neuschwander for their assistance. Furthermore we appreciate the support of the Institute for Applied Research (IAF) Konstanz, especially Prof. Dr.-Ing Horst Werkle and Dipl.-Ing. FH Andreas Burger.

Inhaltsverzeichnis

Institute Profile

Institute for Optical Systems

Brauneggerstraße 55
Building E
3rd and 4th floor
78462 Konstanz

Location Plan

Institute Members

Prof. Dr. Georg Umlauf:

Diploma in computer science from University of Karlsruhe, 1996. Doctoral degree in computer science from University of Karlsruhe, 1999. PostDoc at University of Karlsruhe and University of Florida, Gainesville, USA, 1999-2000. Software development and senior researcher at Tebis AG, Hamburg, 2000-2002. Assistant professor for geometric algorithms at University of Kaiserslautern, 2002-2009. Interim professor for computer graphics at University of Karlsruhe, 2009. Since 2009 professor for computer graphics at University of Applied Sciences Konstanz and head of the computer graphics lab. Since 2010 member of the 'Institute for Optical Systems (IOS)' and 'Institute for Applied Research (IAF)'. Main research interests: Computer graphics, geometric modeling (splines, subdivision), reverse engineering, physical simulations.

Prof. Dr. Matthias Franz:

M.Sc. in Atmospheric Sciences from SUNY at Stony Brook, Diploma in physics from the Eberhard-Karls-Universität, Tübingen and doctoral degree in 1998. Thesis research in visual insect and robot navigation at the MPI for Biological Cybernetics and as a PostDoc at the Australian National University in Canberra. In industry he worked on various aspects of autonomous vision systems. 2002, he returned to the MPI as a group leader in the area of machine learning and computer vision. Since 2007 professor at the University of Applied Sciences in Konstanz and head of cognitive systems lab. Member of 'Institut für Angewandte Forschung (IAF)'. Main research activities in the development of automatically generated vision systems, optimisation and probabilistic modeling, with applications in industrial machine vision, texture analysis and steganalysis.

Prof. Dr. Claus Braxmaier:

Diploma in precision engineering at University of Applied Sciences Furtwangen. Diploma in physics and doctoral degree at the University of Konstanz in the field of fundamental tests of physics. Post-Doc at University of Konstanz. At EADS Astrium GmbH, system responsible for scientific and Earth observation missions for ESA and head of group 'Mission Metrology'. 2005-2013 professor for physics and control theory at the University of Applied Sciences Konstanz. Since 2013 ZARM Deputy Executive Director, Director Space Technology, and department lead SSystem Enabling Technologiesät the DLR-Institute of Space Systems, Bremen. Main research: high resolution optical metrology for industrial and space applications, tests of fundamental physics.

Prof. Dr. Bernd Jödicke:

Study of physics at the University of Kalrsruhe. Doctoral degree at Technical University Hamburg and University Karlsruhe in the field of high frequency technology. After that, industrial work at ABB Baden, Switzerland, as executive director for R&D. Since 1992 professor for applied physics at University of Applied Sciences Konstanz. Member of 'Institute for Applied Research (IAF)', 'Institut für Naturwissenschaften und Mathematik (INM)', 'Institute for Optical Systems (IOS) Konstanz' and 'Deutsche Lichttechnische Gesellschaft'. Head of laboratory for light engineering at HTWG. Main research activities in color and light measurements and color camera systems.

Prof. Dr. Burkhard Lehner:

Diploma in computer science from University of Kaiserslautern, 2004. Doctoral degree in computer science from University of Kaiserslautern, 2008. Software development at Sirona Dental System GmbH, Bensheim, 2008-2013. Since 2013 professor for computer science at University of Applied Sciences Konstanz. Since 2014 member of the 'Institute for Optical Systems (IOS)'. Main interests: software development, computational geometry, optical 3D measurement (especially in dental CAD/CAM).

Prof. Dr. Klaus-Dieter Durst:

Study of physics at the University of Stuttgart, 1986 doctoral degree in the field of magnetism at the Max-Planck-Institute of metal research. Thereafter research center Weissach of the Dr. Ing. h.c. F. Porsche AG, responsible for the central unit 'measurement technologies'. Since 1993 professor for measurement engineering and sensor technology at the University of Applied Sciences Konstanz. Member of 'Institut für Naturwissenschaften und Mathematik' and 'Institute for Optical Systems' Konstanz. Head of laboratories for measurement and sensor technology and production metrology. Currently director of 'Institut für Naturwissenschaften und Mathematik' Konstanz. Activities in the accreditation and surveillance of testing laboratories and inspection bodies.

Prof. Dr. Christian Hettich:

Diploma in physics and doctoral degree (2002) at the University of Konstanz in the fields of quantum-optics and nano-optics. Post-doc at the Eidgenössische Technische Hochschule (ETH) Zürich and at the Niels Bohr Institut in Kopenhagen. Systems engineer for illuminators of semiconductor lithography systems at Carl Zeiss SMT GmbH in Oberkochen. Systems engineer for time-of-flight cameras at ifm automotive GmbH in Kressbronn. Project leader for automated microscopes at Sensovation AG in Radolfzell. Since 2014 professor for physics, feedback control systems and metrology at the department of mechanical engineering of the University of Applied Sciences Konstanz. Member of the 'Institut für Naturwissenschaften und Mathematik' (INM) and the 'Institute for Optical Systems' (IOS). Main research interests: optical 3D measurements, spectroscopy, interferometry, holography, microscopy.

IOS Staff

Professors	Georg Umlauf, director IOS Matthias Franz, associate director IOS Claus Braxmaier Bernd Jödicke Burkhard Lehner Klaus-Dieter Durst Christian Hettich
Officer	Pascal Laube
Academic Staff	Leonard Thießen Martin Miller Mirco Indlekofer Theresa Kocher Dennis Grießer Fabian Freiberg Tobias Birkle Marco Fehrenbach
Postdoc	Thilo Schuldt
PhD Students	Michael Grunwald Martin Schall Pascal Laube Jürgen Keppler

External Fundings and Grants

- Baumer Inspection GmbH, Konstanz: "Farbtexturen in der industriellen Oberflächeninspektion", contract research.
- Baumer Inspection GmbH, Konstanz: Ïnline - Inspektionstechnologie zum Farbabgleich für den digitalen Dekordruck", contract research.
- Australian Research Council, Pattern recognition in animals and machines: using machine learning to reveal cues central to the identification of individuals, Discovery Projects Grant.
- BMBF-Grant "Forschung für die Produktion von morgen"2015, Entwicklung einer innovativen Anlagentechnik zur automatisierten und laserbasierten Reparatur strukturierter Formeinsätze - ToolRep.

Cooperations with Research Institutions and Industry

Academic and Institutional Cooperations

- University of Queensland, Brisbane
- Humboldt-Universität zu Berlin
- ZARM (drop tower), Center of Applied Space Technology and Microgravity, Bremen
- DLR Institut für Raumfahrtsysteme Bremen
- University of Tübingen
- Max-Planck-Institute for Biological Cybernetics, Tübingen
- German Federal Office for Information Security (BSI), Bonn
- Universität Konstanz
- University of California, Davis
- Technische Universität Kaiserslautern
- University of Florida, Gainesville
- Grenoble Institute of Technology
- University of Strasbourg

Industry Cooperations

- Siemens Postal, Parcel & Airport Logistics GmbH, Konstanz
- Sirona GmbH, Bensheim
- EADS Astrium, Immenstaad
- Breuckmann GmbH, Meersburg
- Chromasens GmbH, Konstanz

- Baumer Inspection GmbH, Konstanz
- Procon-System GmbH, Thierstein
- Lightdesign-Solutions GmbH, Dresden
- Knotenpunkt, Wenzel Präzision GmbH, Balingen
- Tebis AG
- Liebherr Aerospace GmbH, Lindenberg
- NTT Data Deutschland GmbH
- Lacuna Solutions GmbH
- EUTECT GmbH
- Siegfried Hofmann GmbH
- ACSYS Lasertechnik GmbH
- Frauenhofer Institute for Production Technology IPT

Theses and Student Projects

PhD Theses

- Bartolomiej Piotr Siwek, Discrete methods for splines and subdivision curves, University of Oslo, 2015.

Master Theses

- Matthias Hermann, Methoden des Deep Learning im Bereich Convolutional Neural Networks, 2015.
- Merlin Blume, 3D Primitive Classification Using Stacked Autoencoders, 2015.
- Patrick Mutter, Automatische Bestimmung von Kontextinformationen aus Bildern mit maschinellem Lernen, 2016.
- Manuel Hieke, Texture recognition based on texton analysis and machine learning algorithms, 2016.
- Martin Miller, Adaptive B-Spline Activation Function for Neural Networks, 2016.
- Mirko Indlekofer, Merkmalsbasierte Objekterkennung mit Methoden des maschinellen Lernens, 2016.
- Theresa Kocher, Defect Detection Using a Parametric Texture Model, 2016.
- Timo Hamm, Vergleich künstlicher neuronaler Netze mit Transformations-Codes, 2016.
- Marco Fehrenbach, Physikalische Simulation für einfache Kamerasysteme, 2016.

Bachelor Theses

- Felix Peter, Optische Selbstlokalisation eines 3D-Druckkopfes im dreidimensionalen Raum, 2015.
- Tarek Schneider, VR-Scan-Simulation Raycasting von Spline-Flächen, 2015.
- Malvin Danhof, VR-Scan-Simulation Raycasting von Dreiecksnetzen, 2015.
- Tobias Ofterdinger, Simulation und Auswertung optischer Filter mit Hilfe eines multispektralen Aufnahmesystems, 2015.
- Benjamin Kugler, Simulation von Objektiveigenschaften anhand radiometrischer Referenzdaten, 2015.
- Robin Mattes, Optische Detektion von Düsenfehlern im Inkjet-Druck, 2015.
- Tobias Birkle, Simulation von Kamerasensoren anhand radiometrischer Referenzdaten, 2015.
- Jens Gansloser, Radiometric calibration of digital cameras using sparse Gaussian processes, 2016.
- Tung Pham, Framelet-Based Image Inpainting, 2016.
- Julian Schneider, Multiskalen Signalverarbeitung von Dreiecksnetzen, 2016.
- Fabian Freiberg, Implementierung eines Convolutional Neural Networks mit CUDA, 2016.
- Dennis Grießer, Geometrische Kalibrierung von Stereokameras, 2016.

Student Projects

- David Simon, Dennis Bleicher, Jürgen Rieg, Point Cloud Label Tool, 2015.
- Simon Kesser, Tobias Keh, Felix Born, Surface Subdivision Tool, 2015.
- Janine Panske, Christian Hümmer, 3d-Rekonstruktion romanisches Portal, 2016.
- Felix Hinderer, Patrick Fiur, Robert Kleinhans, 3d-Fischrekonstruktion, 2016.

Journal Papers

- Schuldt, T., C. Schubert, M. Krutzik, L. Bote, N. Gaaloul, J. Hartwig, H. Ahlers, W. Herr, K. Posso-Trujillo, J. Rudolph, et al., "Design of a dual species atom interferometer for space", Experimental Astronomy, vol. 39, no. 2, pp. 167-206, 2015.

Conference Proceedings

- Denker, K., B. Hamann, and G. Umlauf, Ön-line CAD Reconstruction with Accumulated Means of Local Geometric Properties", Curves and Surfaces, 8th International Conference, Paris 2014: Springer, pp. 181-201, 2015.

- Grunwald, M., J. Müller, M. Schall, P. Laube, G. Umlauf, and M. O. Franz, "Pixel-wise Hybrid Image Registration on Wood Decors", BW-CAR— SINCOM, pp. 24, 2015.

- Schall, M., M. Grunwald, G. Umlauf, and M. O. Franz, "Radiometric calibration of digital cameras using Gaussian processes", SPIE Optics+ Optoelectronics: International Society for Optics and Photonics, 2015.

- Danhof, M., T. Schneider, P. Laube, and G. Umlauf, Ä Virtual-Reality 3d-Laser-Scan Simulation", BW-CAR— SINCOM, pp. 68, 2015.

- Schall, M., M-P. Schambach, and M. O. Franz, Ïmproving gradient-based LSTM training for offline handwriting recognition by careful selection of the optimization method", Conference: BW-CAR Symposium on Information and Communication Systems (SInCom), 12/2016.

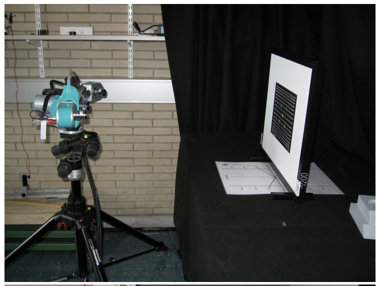
- Schall, M., M-P. Schambach, and M. O. Franz, Ïncreasing robustness of handwriting recognition using character n-gram decoding on large lexica", 12th IAPR International Workshop on Document Analysis Systems: IEEE, 2016.

- Grunwald, M., J. Gansloser, and M. O. Franz, "Radiometric calibration of digital cameras using sparse Gaussian processes", Workshop Farbbildverarbeitung, 2016.

- Laube, P., and G. Umlauf, Ä short survey on recent methods for cage computation", BW-CAR— SINCOM, pp. 37, 2016.

- Grunwald, M., and M. O. Franz, "Wahrnehmungsorientierte optische Inspektion von texturierten Oberflächen", INFORMATIK 2016, Lecture Notes in Informatics (LNI) Gesellschaft für Informatik, vol. 259, pp. 1963-1968, 2016.

Laboratories and Infrastrcture

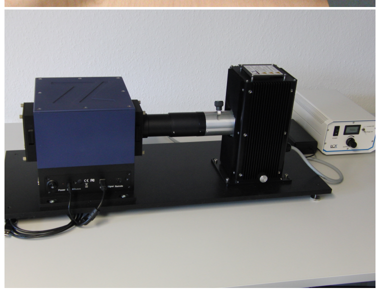


The Institute for Optical Systems laboratories and infrastructure is composed of scientific equipment from the fields of Image Sensing, 2D and 3D Surface Analysis and Reconstruction as well as a powerful computing cluster for machine learning and high performance computing. Our laboratories and infrastructure is located in buildings *E*, *G* and *O* on the university campus.

The laboratory in building G includes equipment with focus on illumination engineering and high precicson stereoscopic surface scanning. The laboratory in buildig E comprises of camera calibration and spectrometric color measurement equipment, multispectral cameras as well as devices for 3D laserscanning, 3D visualization and 3D printing.

The computing cluster located in building *O* consists of GPU as well as CPU-Servers for high performance computing.

Most of the institutes workspaces are located on floors three and four of building *E*.



Research Activities

Reconstruction of Petershausen's Abbey Portal

Christian Hümmer, Janine Panske, and Georg Umlauf

This project is concerned with the 3d reconstruction of the roman portal of the former minster of Petershausen. The only remains of this portal consist of the tympanum, the brow-piece, four capitals and two jamb statues. These artifacts are exhibited at the Badische Landesmuseum, Karlsruhe.

The goal is to digitize these artifacts as good as possible to allow for a subsequent 3d reconstruction. To this aim various scanning technologies were used. Each device yields different types of data at different resolutions and different noise signatures. However, finally only one merged, de-noised, and compressed 3d model of the artifacts is required.

Introduction

The abbey in Pertershausen was founded in the 10th century. The minster was teared down 1831 and the remains of the portal were shipped to Eberstein. The respective artifacts of the portal of the minster were made between 1173 and 1180, see Figure 1 (left). The two jamb statues portray pope St. Gregor I. (590 – 604) and St. Gebhard II., bishop of Konstanz and founder of the abbey (979 – 995) They are made of yellow-gray colored sandstone, which was exposed to various environmental stresses like weather-caused erosion, a fire, and multiple dismounting, transportation, and re-mounting procedures. Thus, the remains are broken, disconnected, and details are eroded. The height of the complete portal is up to four meters and weights several tons.

These remains are in a permanent exhibition in the Badischen Landesmuseum in Karlsruhe, see Figure 1 (left). However, this permanent exhibition is timbered in a closed and tight wooden shed. This means, that the accessibility is very limited and the lighting situation is extremely difficult, see Figure 2.

Abbildung 1: Con-temporal painting of the minster's portal of Petershausen (left) and current display in the Badischen Landesmuseum, Karsruhe, (right) [1]. Courtesy: Badisches Landesmuseum.

Abbildung 2: During the scanning.

Project description

In October 2016 a team of several students and staff from HTWG as well as from our industry partner FARO meet in Karlsruhe for the digitization of the artifacts. For the digitization of the artifacts various scanning techniques and devices have been used:

- FARO laser-line probe, a manual laser scanner mounted on a measuring arm. Reconstruction results are shown in Figure 3.
- FARO Freestyle 3D, a manual structured-light scanner. Reconstruction results are shown in Figure 4.
- Structure Sensor, a manual infrared structured-light scanner. This techniques yield results that could not be used for a detailed reconstruction.
- Photogrammetry based on digital camera images and a subsequent Agisoft-

reconstruction. Reconstruction results are shown in Figure 5.

The complete reconstruction process consists of four stages

1. digitization,
2. point cloud processing (registration, subsampling, de-noising, etc.),
3. mesh generation (Poison, ball-pivoting, etc.), and
4. post processing (texturing, etc.).

The final 3d models of the artifacts were provided to the Faculty of Architecture of HTWG, to be used in the design of a complete physical reconstruction of the portal at its original location, and to a toy manufacturer to reproduce the two jamb statues at a size of ca. 30 cm for the Verein Petershauser Orgelkultur as artwork presents [2].

Abbildung 3: Laser-scanner reconstruction of the jamb statue of pope Gregor I.

Abbildung 4: Structured-light scanner reconstruction of the tympanum and the brow-piece.

Abbildung 5: Photogrammetric reconstruction of the portal.

Literaturverzeichnis

[1] https://www.leo-bw.de/en_US/web/guest/detail-gis/-/Detail/details/DOKUMENT/blm_museumsobjekte/8A4B2C47F3AF49488B6F0F6BE6EF871D/Kirchenportal+Kirchenportal+sog+%22Petershausener+Portal%22

[2] https://www.petershauser-portal.de/

Detection of flowmarks on laminated decor using texture features

M. O. Franz, H.-P. Diehl[1]

The production process of laminated decors is subject to a large variety of perturbations some of which are very hard to detect by machine vision algorithms. Among the most notorious are flowmarks that are caused by processing errors during lamination. Flowmarks are only visible in certain lighting and viewing conditions, and only show up with a very small texture contrast. In this project, we developed a new detection algorithm for flowmarks that is based on higher-order texture features in combination with support vector classifiers.

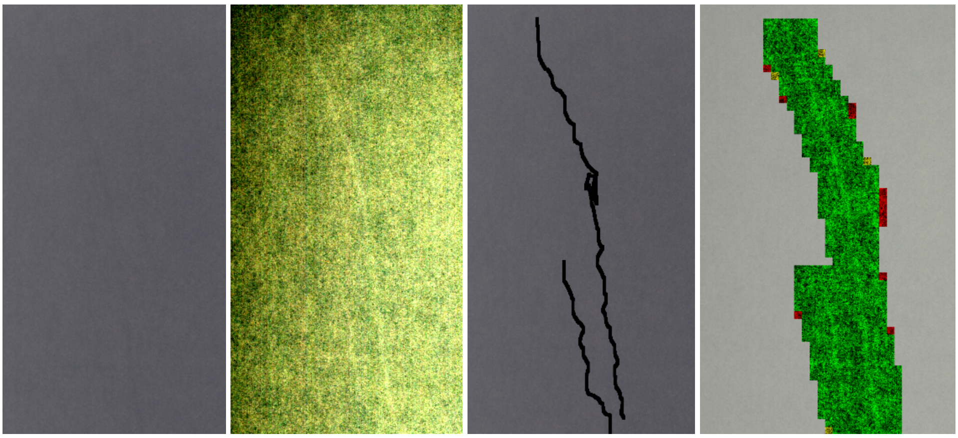


Abbildung 6: *Left:* Example of flowmarks on a laminated decor shown in the original contrast; *middle left:* Contrast maximized image of the flowmarks; *middle right:* Hand labeling for training and testing the detection system; *right:* detector output - detected regions that contain flowmarks are marked in green, erroneous detections in red.

Flowmarks usually appear as low-contrast linear structures that are very hard to detect in a textured background (see Fig. 1, left). However, to the human eye these structures are still visible under certain viewing angles and possibly disturbing. From the point of view of machine vision, flowmarks are particularly hard to detect as they show themselves only by small alignments of texture elements, not by different contrasts or colors. In terms of statistics, this means that flowmarks can only be detected by analyzing higher-order pixel dependencies, i.e. dependences between three or more pixels. We therefore used higher-order texture features that were previously only applied to problems in steganalysis[2]. For training our detector, we labelled flow marks by hand and subsequently divided the images in overlapping regions (see Fig. 1, middle right). The texture in each region was characterized by its higher-order features which comprised the input to a support vector machine. The training of the detector required about 5000 image regions before a satisfactory detection rate could be achieved. The trained detector has a very low false alarm rate while correctly marking the majority of all image regions contaminated by flowmarks (see Fig. 1, right).

[1] Baumer Inspection GmbH, Lohnerhofstrasse 6 78467 Konstanz

[2] Pevny, T., et al.: Steganalysis by subtractive pixel adjacency matrix. IEEE Trans. Info. Forensics and Security, 5(2), 215-224, 2010.

LabelImage Tools

Sonja Futterknecht, Tobias Birkle, Benjamin Kugler, Robin Mattes and Michael Grunwald

Machine learning needs a lot of training data to detect objects in images. To generate the training data, we developed a software soultion to label and classify fishes in images. This project was in cooperation with the Global Change Institute of the University of Queensland.

Introduction

Regarding the number of fishes and their diversity it is possible to investigate the consequences of the climate change in coral reefs. This is the project of the Global Change Institute (GCI) at the University of Queensland under the head of Prof. Dr. Ove Hoegh-Guldberg. In order to specify the number of fishes experts dive through the coral reefs and document the fish population. Another project at the GCI is the *XL Catlin Seaview Survey*, in which coral reefs are photographed using underwater cameras, as shown in figure 7. The images of the reefs are published on the website *globalreefrecord.org*. The *FishNet* project, under the management of Dr. Ulrike Siebeck from the GCI wants to use the underwater photos of the coral reefs from the *XL Catlin* project to identify fishes and their number in the images. To detect the fishes a software soultion will be developed using methods of machine learning. In the first step the fishes have to be labeled and classified in the photos manually to extract their features. For this part of the software solution a software called *LabelImage* is developed to label and classify the fishes. These labeled fishes are the training images for the machine learning. Based on the training images the features of the fishes will be learned and it will be able to find these features in not yet labeled images. The manual labeling and classification of the fishes will be done by the experts of the *FishNet* team.

Current Work

The software was developed with Python 3.4 and PyQt 5. With this application a rectangle can be drawn around the fishes with one mouse click to label and classify them (see fig. 8). The coordinates of the labeled area and the corresponding fish species will be saved in an XML document. In a separate extraction module of the software the labeled fishes will be extracted from the original image with the help of the XML document. The extracted patches will be saved as single images. The patches can be used as training images for the machine learning. They form the basis for the automatic identification of the fishes. The determination of the species diversity and the number of the fishes works the better the more data is available for the software solution. To submit many fishes and thus training data to the soft-

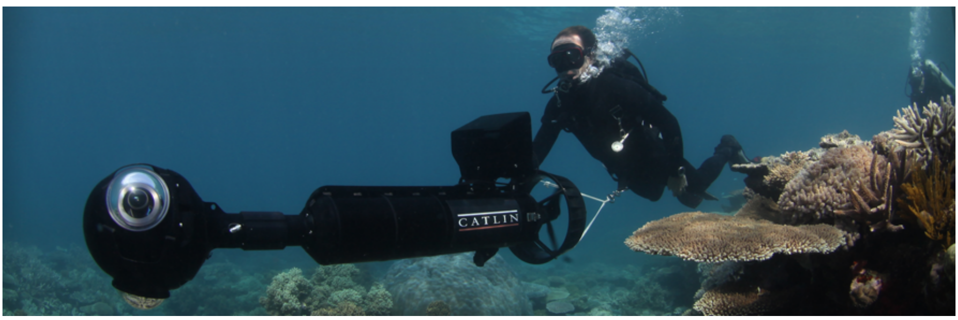


Abbildung 7: *XL Catlin Seaview Survey* expert diving through coral reef. (globalreefrecord.org)

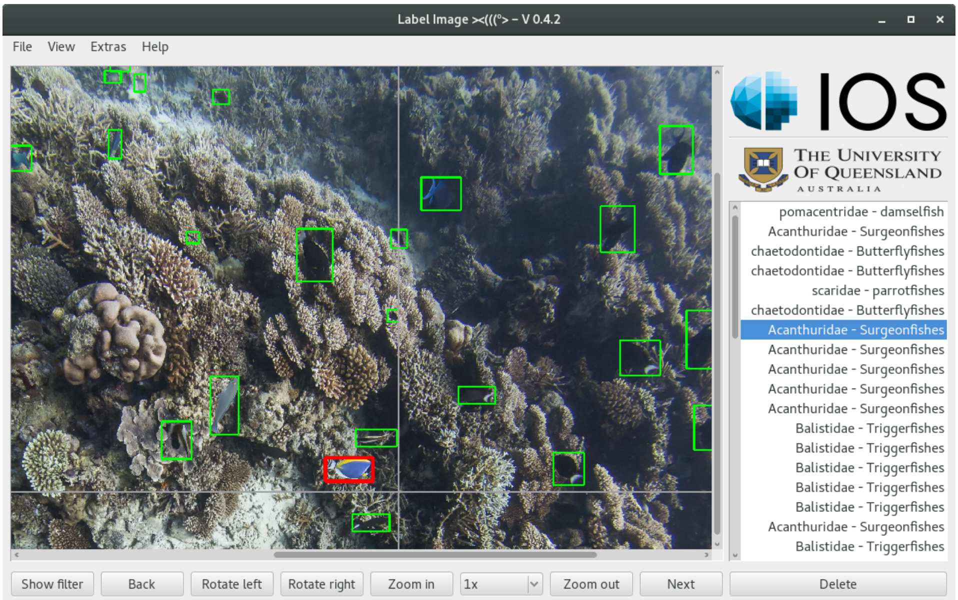


Abbildung 8: *LabelImage* software with labeled fishes.

ware solution the fishes are not only labeled by the scientists of the GCI but also by interested users. Therefore in the second part of the team project a web based software solution was developed using HTML5 and the JavaScript framework jQuery which allows the labeling and classification of the fishes. The software solution is embedded in an online course at the Massive Open Online Course platform *edx.org* and developed additionally as a standalone website. The labeled data will be saved in the same XML format as in the *LabelImage* software. Thus the extraction module of the software can be reused for the labeled patches in the web based software. The students will be evaluated by the number of labeled fishes, with an especially developed evaluation software of the team project. For the evaluation a reference image is used in which the fishes are already identified by the experts of the GCI. If the results are satisfying it can be assumed that the following images will be labeled equally well. These data contributes to the machine learning for the automatic identification of the fishes. The online course will be repeated every eight weeks to collect training images for the *FishNet* project. During the first two weeks after the activation of the course at the platform edx.org over 300 participants took part in the course. Thereby over 13500 fishes were found in 600 images. In comparison to that, the experts of the *FishNet* team labeled 5358 fishes in four weeks. The experts first selected eleven fish species that the participants could then associate with the fishes in the pictures. In average the participants labeled 80% of the pictured fishes and associated 25% to a fish species.

Conclusion and Outlook

In summary it can be said that fishes can be labeled and classified in images with the *LabelImage* software solution by the scientists and with the web based software solution by other participants. With the evaluation software participants from the online course and the standalone website can be evaluated and the labeled fishes can be processed for the machine learning. Using these images the software can be trained and the classification of the fishes will be optimized. The software solution for identifying the fishes should facilitate the work of the scientists from the GCI. They do not have to classify the fishes manually anymore. With the underwater images from the *XL Catlin* project and the software solution the changes of the fish population in the coral reefs all over the world can be captured faster. The machine learning can be also used for detecting other objects in images, if there are enough training images available. Therefore any images can be loaded to the *LabelImage* software and then the desired objects can be labeled and classified.

Framelet-based image inpainting

Tung Pham, Pascal Laube and Georg Umlauf

Image inpainting is an important problem in the domain of image processing. Tight wavelet frames which are a specialization of well knot wavelets have recently been successfully applied to the problem of image inpainting [1]. In this project we evaluated the performance of framelet-based image inpainting on classic test images.

Introduction

Filling missing regions in images is a prominent topic of image processing. Image inpainting may help in image transmission, image coding or restoration. Many popular methods use the principle of filling a missing region by sampling from a "healthyßurrounding region. Results should look reasonable to a human observer. In this project we use wavelet tight frames or framelets for image inpainting. By first transforming the image to the framelet domain and applying thresholding in this domain one can fill in previously unknown information.

Method

Let $\langle \cdot, \cdot \rangle$ and $\| \cdot \|^2$ be the dot product and norm of $L_2(\mathbb{R})$. A wavelet-system $X(\Psi)$ is a family of scaled and translated wavelets of finite set $\Psi \subset L_2$:

$$X(\Psi) = \{\psi_{s,u} : \psi \in \Psi, \quad s, u \in \mathbb{Z}\}$$

Elements ψ of Ψ are called generators. A system $X(\Psi)$ is called *Bessel-system*, if there exists an upper bound

$$\sum_{g \in X(\Psi)} |\langle f, g \rangle|^2 \leq B \| f \|^2, \forall f \in L_2(\mathbb{R}), B > 0.$$

A Bessel-system $X(\Psi)$ is called *frame*, if there exists a lower bound

$$\sum_{g \in X(\Psi)} |\langle f, g \rangle|^2 \geq A \| f \|^2, \forall f \in L_2(\mathbb{R}), A > 0.$$

If $A = B$ this is called a *tight frame*:

$$\sum_{g \in X(\Psi)} |\langle f, g \rangle|^2 = \| f \|^2, \forall f \in L_2(\mathbb{R}), A = B = 1.$$

If $X(\Psi)$ is a tight frame, generators $\psi \in \Psi$ are called *tight framelets*. Compared to orthonormal wavelets, tight framelets are redundant. This redundancy can be exploited for the purpose of inpainting.
The inpainting process consists of four steps:

1. By applying an analysis operator the image is first transformed to the framelet domain where it is represented by its framelet coefficients.

2. To propagate information into the defective region framelet coefficients are tresholded.

3. By then synthesizing back to the image domain with thresholded coefficients missing information is filled.

4. For smooth transition this process is repeated until converge criteria are met.

Hard thresholding is defined as

$$T(c_n), = \begin{cases} 0, & |c_n| \leq T, \\ c_n, & \text{sonst} \end{cases}$$

with framelet coefficients c_n and threshold T. By applying this hard thresold small coefficients are completely removed. This also suppresses static noise. An open question is the selection of T. Large T results in missing detail while small T leads to insignificant changes.

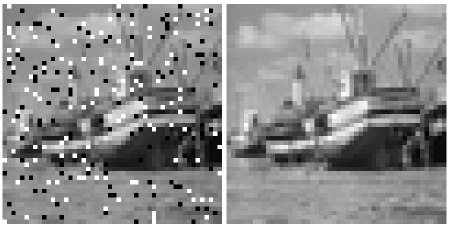
For *soft thresholding* every coefficient receives an individual threshold value.

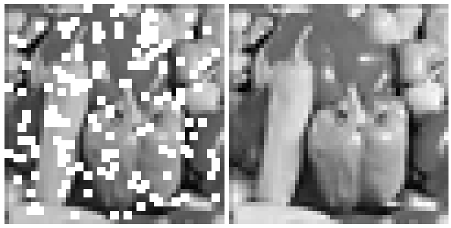
$$t_{\lambda_i}(c_n) = \begin{cases} 0, & |c_n| \leq \lambda_i, \\ sgn(c_n)(|c_n| - \lambda_i), & |c_n| > \lambda_i \end{cases}$$

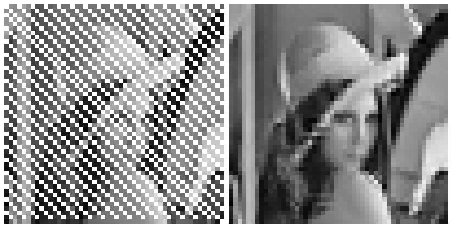
While low-pass coefficients are not manipulated in hard-thresholding they are in soft thresholding. This can help the inpainting process.

Results

Following we show some results of framelet-based inpainting.







Literaturverzeichnis

[1] Cai, Jian-Feng and Chan, Raymond H and Shen, Zuowei: A framelet-based image inpainting algorithm. Applied and Computational Harmonic Analysis, 24, 131–149, 2008.

Vector Machines for Knot Placement in B-Spline Curve and Surface Approximation

Pascal Laube, Georg Umlauf, Matthias Franz

Knot placement for curve and surface approximation is an open problem. Selecting knot values to receive good approximation results is a challenging task. Proposed approaches range from parametric averaging to genetic algorithms. We propose the use of Support Vector Machines (SVMs) for finding suitable knot vectors in B-spline curve approximation. The SVMs are trained to be able to distinguish between locations along the curve that are well or not well suited as knots in the parametric domain. This score is based on different geometric features of a parameters corresponding point in the point cloud. A score weighted averaging technique is used to produce the final knot vector. We further propose a method to use the score weighted averaging technique for T-spline surface approximation.

Finding a good knot vector for curve and surface approximation is a common problem in many applications like e.g. Reverse Engineering (RE). In RE one starts with a sampled physical object represented by a point cloud. For this point cloud a corresponding CAD representation must be recover. After steps such as pre-processing and segmentation, a B-spline approximation of the point cloud or parts of it is computed. For this approximation the data points, parameter values for the data points, the B-spline degree, and an appropriate knot vector are required. Then the goal is to generate knot vectors that lead to approximations with as little deviation from the data set, for as few knots as possible. Usually a certain threshold has to be satisfied to assure a sufficient approximation quality. Since the number of control points is determined by the number of knots, the number of knots is especially important if the resulting curves and surfaces will be further processed by a human operator.

Because there are many unknowns in the approximation process, we propose knot vector computation using machine learning methods. We train Support Vector Machines (SVMs) to distinguish between locations along a curve that are well or not well suited as knots in the parametric domain. For training and classification we use a feature vector that concentrates information about a point clouds geometric structure. Consider a set of ordered points in two dimensional space. We define a good knot vector as the one that generates little deviation to the data set while having a small number of knots. Finding the ideal knot vector is impossible. A computationally very expensive way is to exhaustively search for appropriate knot value combinations in the set of parameters. Due to computational cost this is not feasible for practical application. Since training data has to be generated only once the cost of exhaustive search in our case is negligible. After training the SVMs are able to assign a score to each point of a point cloud based on its local point neighborhood.

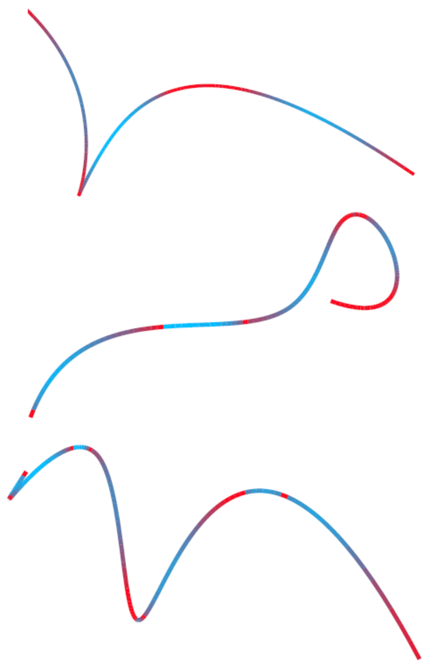


Abbildung 9: Three exemplar curves colored by score from high score (red) to low score (blue).

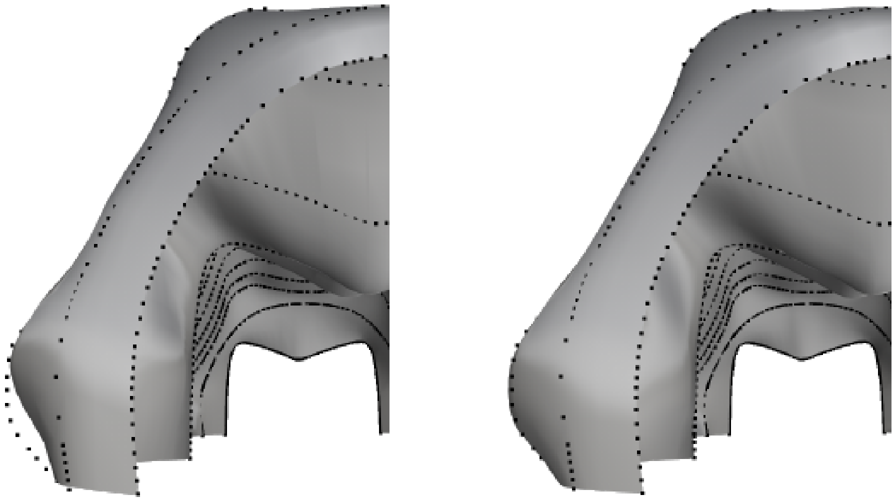


Abbildung 10: T-spline surface approximation by NKTP on the left and by SWKP on the right together with the original point data.

An score-weighted averaging yields then the final knot vector. This concept for B-Spline curve approximation is then applied for T-spline surface approximation. B-spline surface interpolation by surface skinning or lofting is the process of interpolation by contour curves (CCs). For an ordered set of points in which each point belongs to a serial contour along the surface one has to first interpolate points along CCs. The compatibility of the CCs needs to be ensured by degree elevation to a common degree. By interpolating the CCs one receives the lofted surface. Approximation of CCs results in an approximate lofted surface. Unordered sets of points can be converted to ordered sets for surface lofting by point cloud slicing. We propose an approach for B-spline curve and T-spline surface approximation. We train SVMs to decide on the suitability of pre-computed parameter values of a point cloud as values for a knot vector. Based on this suitability measure, or score, we perform a weighted averaging of parameter values to create the knot vector. This concept is further adapted for T-spline surface approximation by lofting. The proposed method produces approximations of good quality. This is especially the case for approximation with a small number of knots. We could show that SVMs are able to learn characteristics of positions along a curve where knot placement has positive impact on approximation quality. The approximation by lofted T-spline surfaces is able to preserve initial knot vectors of CCs and needs a fraction of the knots a comparable B-spline surface would need to satisfy a certain threshold. For future works we plan to investigate the applicability of machine learning to other steps of the RE process like point cloud segmentation. We further would like to apply deep learning techniques that render manual feature selection obsolete like e.g. neural networks with an auto-encoder layer.

Morphed 3d fish models using multi-perspective 2d fish images

Leonard Thiessen, Pascal Laube, Ulrike Siebeck and Georg Umlauf

In this work we present a method to generate 3d fish models based on 2d fish images. First 3d and 2d contours are extracted. Then correspondeces are computed and the shape of the 3d model is morphed, using a spring-mass system, to the 2d reference contour. 3d vertex coordinates are interpolated to yield the final 3d fish model.

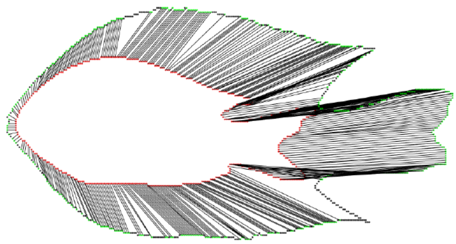
Introduction

Based on the long-standing research cooperation of the HTWG Konstanz, in particular the Institute of Optical Systems, and the Faculty of Biology of the University of Brisbane, joint research projects take place regularly. The aim of the research project "Fishmorphïs to generate lifelike reconstructions based on image data, 3D data and sample models of various fish to be able to realistically map the ichthyological conditions at the Great Barrier Reef on the endemic species there. In order to be able to receive the partially species-protected fish, the fish must be measured alive and then reset. The 3d reconstructions will be used in further projects to investigate the behaviour of fish. In recent years the research in the field of behavior and perception for fishes has been insightful. Is has been discovered, that particular species of reef fish are able to detect simple and complex shapes as shown in [3], and that these fishes can also use ultraviolet light to detect other fishes [2]. And as recently shown in [4] the archerfish has the capability to discriminate human faces. This is remarkable, because this ability is attributed to the cortex, but a fish has no cortex.

Another interesting research purpose for 3d reconstructions is the perception of fishes, especially the way how a fish perceives other fishes and what the crucial point are, where a fish recognizes another fish as a predator or another harmless fish. Discovering this behavior requires experiments with fishes. Here one problem is that some of the fishes are endangered, but they have proven as relevant for this kind of research. This could be avoided by usage of an animation for such experiments as proposed in [5]. The usage of computer animations or 3d models in such a research field is already ongoing. In [4] 3d models of human faces were used for this experiment. So 3d models of fishes could support the current research purposes. The problem remains, that there are no accurate fish models.

Method

The focus of this work was to generate new 3d models of fishes based solely on texture images. Therefore, a 3d prototype model for each desired fish species or family is required and the images should show the respective fish from different perspectives, since one image can only model one perspective. Out of these different perspective we extract the fishes shape (or contour).



A 3d fish prototype model will then be adapted (or morphed) to match the fishes reference shapes. Shape matching is applied to discover correspondences between the contour of the 3d model and the 2d reference contours which allows to adapt the contour of the model to the reference's contour. Morphing the 3d model is done by applying a mass-spring system in image space and then interpolating 3d vertices with regard to the morphed pixels. The mass-spring system determines where which parts of the prototype model should be moved to, to generate a 3d model that looks like the reference fish. This process has to be repeated several times for each fish perspective to get an accurate 3d model in the end.

Conclusion

We could show, that the approach of a texture-based deformation of 3d models, is able to produce accurate 3d fish models. The 2d/3d morphing achieved results which can be used for research purposes in the field of behavioural research of fishes. Obviously the results are heavily depending on the input data and especially contour extraction and contour matching. For future work we plan to introduce a 3d mass-spring system which is able to preserve the 3d models volume to counteract volume degeneration.

Literaturverzeichnis

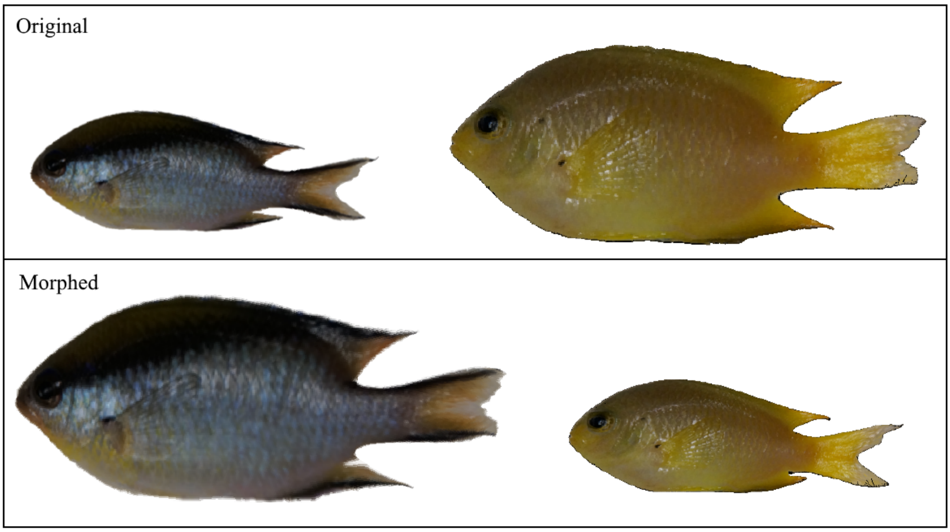
[1] Cai, Jian-Feng and Chan, Raymond H and Shen, Zuowei: A framelet-based image inpainting algorithm. Applied and Computational Harmonic Analysis, 24, 131–149, 2008.

[2] Siebeck, U. E., Parker, A. N., Sprenger, D., Maethger, L. M., and Wallis, G.: A species of reef fish that uses ultraviolet patterns for covert face recognition. Current Biology, 20(5), 407-410, 2010.

[3] Siebeck, U. E., Litherland, L., and Wallis, G. M.: Shape learning and discrimination in reef fish. Journal of Experimental Biology, 212(13), 2113-2119, 2009.

[4] Newport, C., Wallis, G., Reshitnyk, Y., and Siebeck, U. E.: Discrimination of human faces by archerfish (Toxotes chatareus). Scientific reports, 6, 27523, 2016.

[5] Woo, K. L., and Rieucau, G.: From dummies to animations: a review of computer-animated stimuli used in animal behavior studies. Behavioral Ecology and Sociobiology, 65(9), 1671, 2011.



Evaluation of Features for SVM-based Classification of Geometric Primitives

Pascal Laube, Matthias O. Franz and Georg Umlauf

In the reverse engineering process one has to classify parts of point clouds with the correct type of geometric primitive. Features based on different geometric properties like point relations, normals, and curvature information can be used to train classifiers like Support Vector Machines (SVM). These geometric features are estimated in the local neighborhood of a point of the point cloud. The multitude of different features makes an in-depth comparison necessary. In this work we evaluate 23 well known and some new features for the classification of geometric primitives in point clouds. Their performance is evaluated on SVMs when used to classify geometric primitives in simulated and real laser scanned point clouds.

Introduction

Reverse engineering (RE) of scanned 3d objects is an active field of research. The number of 3d scanning devices, e.g. mobile phones, is ever increasing. The result of a scanning process is an unstructured 3d point cloud. To further process this point cloud its CAD representation has to be recovered.
RE consists of three main steps: pre-processing, segmentation, and fitting. Pre-processing includes e.g. subsampling or filtering the point cloud. The segmentation step yields patches in a point cloud that belong to the same parametric CAD model. Often measures of similarity like smoothness or color are used to to segment point clouds. These segmentation methods lack information about the underlying parametric model. For RE the resulting segments have to be classified as either geometric primitives (planes, spheres, cylinders, etc.) or free form surfaces. Depending on the classification result, in the fitting step a suitable CAD model is computed to approximate the point cloud patch.

In this work we concentrate on the classification of geometric primitives prior to the fitting process. The considered geometric primitives are cones, planes, cylinders, ellipsoids, spheres, and tori.

Method

A common approach in RE is to use a RANSAC based approaches. Different parametric models are iteratively fitted and the one yielding the smallest error is assumed to be correct. RANSAC based algorithms have two problems:

- Iterative fitting for each possible primitive type is expensive, if there is no initial knowledge of the type.
- Noisy point clouds might be approximated by the wrong primitive, because it yields the smallest error.

A different approach is to use local differential geometric properties and thresholds for classification. This method requires user defined thresholds that may vary with different scanner types. Recently, machine learning approaches have superseded manual threshold definition. In [1] Support Vector Machines (SVM) together with curvature features are used to detect some types of geometric primitives. In [2] a set of feature histograms is used and their performance for k-nearest neighbor, k-means, and SVM classifiers are compared. In these papers only a small set of features is used. It is common to use all available features for classification even if some of them might weaken classification performance. We present an in-depth comparison of different feature descriptors. We evaluate the performance of 19 individual features and their combinations. Our results give an indication on which features are meaningful for the RE process. We also evaluate the impact of homogenization on classification results.

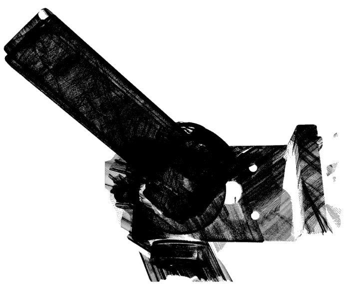
Results and Conclusion

We present the evaluation of normal-, point-, and curvature-based features for primitive recognition in point clouds using support vector machines.

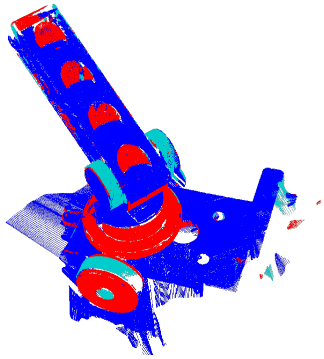
Based on simulated scans we compare the performance of different features and feature combinations. Resulting classifiers were applied to real scans with and without homogenizing the density. Results of curvature-based features did not meet our expectations. Our results can be used to optimize the feature selection for the classification task at hand. We could show that homogenization improves the classification performance on real scanned point clouds.

For future work we intend to use unsupervised learning methods, e.g. auto encoders, for feature engineering. To generate simulated scans that match real scans as close as possible is another aspect we plan to investigate.

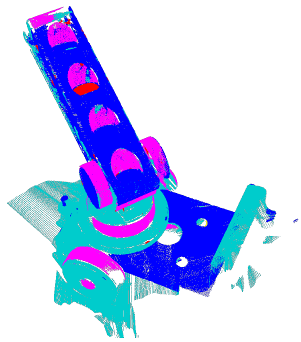
(a) Wooden toy firetruck.



(b) Point cloud of scanned firetruck.



(c) Point cloud of firetruck colored by primitive class.



(d) Point cloud of firetruck colored by primitive class without homogenization.

Abbildung 11: A real scan showing a part of a wooden toy firetruck colored by primitive class using the simple surflet combinations feature: cones(red), planes (blue), cylinders (magenta), spheres (cyan), ellipsoids (green), tori (yellow).

Literaturverzeichnis

[1] Arbeiter, G. and Fuchs, S. and Bormann, R. and Fischer, J. and Verl, A. . Ëvaluation of 3D feature descriptors for classification of surface geometries in point clouds". Inter. Conf. on Intelligent Robots and Systems, 2012.

[2] Rusu, R.B. and Marton, Z.C. and Blodow, N. and Beetz, M. . "Learning informative point classes for the acquisition of object model maps". Intl. Conf. on Control, Automation, Robotics and Vision, 2008.

Analysis of 3d-Convolutional Autoencoders for Symmetry Operations

Felix Hinderer, Pascal Laube and Georg Umlauf

Recently neural networks have become an important tool in CAD-related tasks such as reconstruction. They have been successfully applied for classification of 3d-scanned objects. While they often outperform classic methods it is not easy to understand their intrinsic representation of geometry. Based on proceeding works on primitive classification (see Figure 12) we suspect that an important geometric property learned by the neural networks are object symmetries. In this work we analyse the learned representation by neural networks in trying to recover information about the importance on symmetries.

Method

To analyse the representation learned by neural networks we propose to evaluate the intermediate representation learned by so called autoencoders. The trained neural networks are evaluated using methods for visualization and neuron correlation. First we will shortly introduce the volumetric 3d representation used for training then we will introduce the concept of autoencoders.

Voxelization Voxels are the 3d counterpart to pixels. They are used to transform 3d models into raster graphics. Furthermore, the required memory, when copared to original point clouds, is lower. We started with $52,000$ of ShapeNet [1] models. The procedure of voxelization is the following:

- First, a so-called bounding box is placed around the model. A bounding box is usually the smallest box that contains the object completely. However, since we convert all our data into 3d cubes, our bounding box must be cube-shaped, otherwise the aspect ratios will be lost.

- Now that we have a cube that completely contains our object, we now divide it by the same value on all axes, so if we want to have a 32x32x32 voxel image through 32.

- Now we iterate over all points and look in which voxel, or partial cube, they are and increase its value by one.

- Finally, we transform the resulting matrix into a sparse matrix. This means that we only notice values that are not equal to zero, which leads to further memory savings.

Autoencoder Classic neural networks need classified training data. The big advantage of autoencoders is that they don't need labelled training data. Another feature of the auto encoders is the hourglass-shaped structure of the network. The input and output layers are both the same size, while the hidden layers in between are smaller. This forces the autoencoder to learn distinctive features in the first step and in the second step to reconstruct the input as exactly as possible from those features.
Autoencoders often are symmetrically, i.e. the layers responsible for decoding occur in reverse order to the encode layers. However, this not always is the case, for example we trained asymmetric auto-encoders, i.e. unequal numbers of encode and decode layers, with good reconstruction results.
In this work a special type of autoencoder was trained, namely a convolutional auto-encoder. A Convolutional Autoencoder can, similar to a CNN, only consist of Convolutional Layers, but it can also contain other layers, such as one or more Fully Connected Layers.

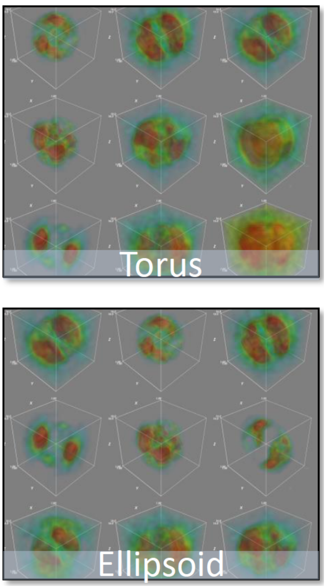
Experimental Results

In our experiments our assumption that symmetries play an important role for neural networks trained on 3d data could not be confirmed. We came to the conclusion that there are several reasons why no symmetry classes, such as from could be isolated:

Data set: In the data set used the symmetries are very unequally represented. While a large part of the data set consists of mirror-symmetric models, there are hardly any n-fold rotational symmetries

and no inversion symmetries. In retrospect, the data set was still supplemented with synthetic data, but this did not affect the results.
Size: The autoencoder can also simply be the wrong size, so that the autoencoder copes well with a linear combination of filters, and it does not need to learn more intelligent representations, such as symmetries.

For future work we plan to apply sparse auto-encoders and a limited data set.



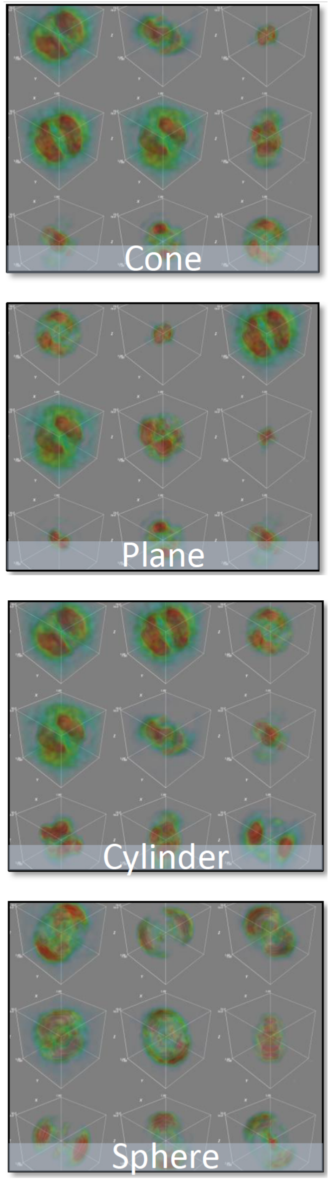


Abbildung 12: Different filter kernels of a neural network trained to classify geometric primitives [2].

Literaturverzeichnis

[1] https://www.shapenet.org/

[2] Blume, Merlin. "3D Primitive Classification using Stacked Autoencoders". University of Konstanz, 2015.

GPGPU and Virtualization for Deep Learning

Martin Schall, Michael Grunwald, Pascal Laube, Matthias O. Franz and Georg Umlauf

Deep Learning has seen some remarkable successes. These include playing the game of Go on a human-master level [1], playing Atari games [2] (both without explicit learning of the games' rules) or the identification of geographic locations from vacation images [3]. Deep Learning is based on the ability to build and train large neural network models using large amounts of data, which is made possible by the introduction of fast, reliable and cheap high-performance computing hardware. Nvidia GPGPU cards and their CUDA API are the de-facto standard for such applications and are widely supported in Deep Learning software libraries. The Institute for Optical Systems is working on providing the needed High Performance Computing platform to provide GPGPU capabilities for its increasing number of scientific projects.

Concept

Experience within the Institute for Optical Systems has shown that a central and powerful computation server is a good approach to making necessary processing power available. This experience combined with the growing number of scientific projects and people at the Institute for Optical Systems has lead to our belief that a reliable, fast and extendable High Performance Computing platform should be based on fast servers and virtualization technology. This allows both the flexible allocation of processing hardware to multiple projects while maintaining clear separation, as well as further extension of the platform to match the requirements of future challenges.

The computing platform is thus based on the open-source and reliable Ubuntu Linux server operating system, KVM virtualization and OpenNebula cloud software. KVM virtualization allows allocation of GPGPU cards to specific virtual machines with near-native performance and has been successfully applied to high performance computation problems by scientific and industrial enterprises. OpenNebula is a cloud-software for managing clusters of servers and virtual machines running on them. It allows the easy and centralized managment of the virtual machines in use, enabling hazzle-free time sharing of the GPGPU hardware between multiple projects.

The overall architecture of the new High Performance Computing platform is shown in abstract form in Figure 13. This system is implemented and in use for research projects at the Institute for Optical Systems.

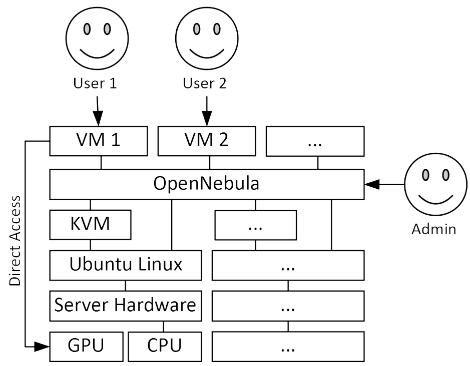


Abbildung 13: Concept of the KVM- and OpenNebula-based High Performance Computing platform.

Hardware

Modern computing hardware has been acquired for use in this High Performance Computing platform at the Institute for Optical Systems. It consists of a 19 inch rack server located in building O of the University of Applied Sciences Konstanz. It contains 32 Intel Xeon CPU cores, running at 2.4GHz as well as 128GB main memory. GPGPU hardware in this system consists of one Nvidia Tesla K20 and one Nvidia Tesla P100. Operating system is a Ubuntu Linux LTS server edition.

This system serves as the main node of the OpenNebula instance and thus the backbone of the High Performance Computing platform.

Deep Learning

Nvidia GPGPU cards and their CUDA API are the de-facto standard for scientific Deep Learning applications. Deep Learning libraries that build on the CUDA API are e.g. Tensorflow and PyTorch. Those libraries are both open-source and in use e.g. at Google and Facebook. The use of modern GPGPU hardware and open-source Deep Learning libraries allows the Institue for Optical Systems to incorporate the most recent scientific advances in the field into its own research projects.

Conclusion

The described High Performance Computing platform has been sucessfully introduced at the Institute for Optical Systems, providing reliable and fast access to GPGPU hardware to enable Deep Learning research. Computing hardware is manufactured by Intel (CPU) and Nvidia (GPU) in the current system. The platform is based on Ubuntu Linux, KVM and OpenNebula and allows easy management and deployment of multiple virtual machines. Strict separation between different projects as well as customization of the virtual machines to the projects' needs is thus maintained.

This new platform has proven to be a valuable asset to the Institue for Optical Systems. Not only current research projects but surely also future ones will rely on this fast and reliable platform.

Literaturverzeichnis

[1] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[2] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with Deep Reinforcement Learning. pages 1–9, 2013.

[3] Tobias Weyand, Ilya Kostrikov, and James Philbin. PlaNet - Photo Geolocation with Convolutional Neural Networks. Technical report, 2016.

Student project - Money Checker

Dennis Grießer, Fabian Freiberg and Michael Grunwald

Comparing image patches is a task of fundamental importance in the field of computer vision. To acquire a measurement of similarity it is necessary to make use of handcrafted features. With the aid of machine learning it is possible to learn a general similarity function for comparing image patches directly from image data. This approach was based on the work of Zagoruyko et al [1] who opted for a CNN-based model.

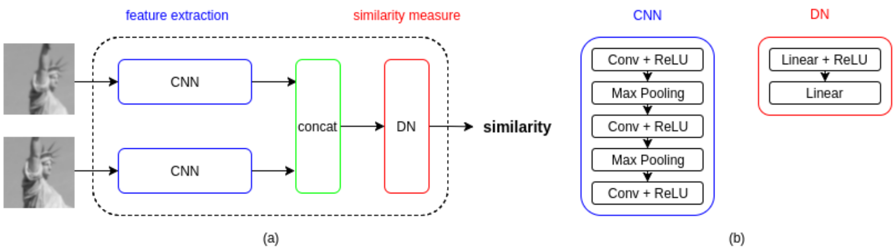


Abbildung 14: (a) Basic structure of our pseudo-siamese network. Two branches with an identical structure consisting of convolutional and pooling layers serving as feature extraction. Both CNN outputs are being concatenated within a merge layer and then being input into the decision network which serves as similarity measure. (b) Detailed structure of CNN and DN, based on [1].

Introduction

The ability to compare image patches has been the basis of many approaches to core computer vision problems, including object, texture and scene categorization. Thus far manually-designed feature descriptors, such as SIFT [3], had to be used, to determine whether or not two patches correspond to each other.

Zagoruyko et al [1] introduced multiple approaches to directly learn such needed features from image data. They studied three neural network architectures, namely 2-channel, siamese and pseudo-siamese.

Alongside the work of Kugler [2] who partially studied siamese neural networks, we focused on the pseudo-siamese architecture. Our goal was to extend ummon such that it is possible to build pseudo-siamese networks with said library and thereby be able to compare image patches via convolutional neural networks.

Architectures

Siamese networks consist of two branches that share exactly the same architecture and the same set of parameters. Each branch takes one of the two image patches as input and then applies a series of layers. Both branch outputs are then concatenated and given to a top network.

The siamese network's branches can be considered as descriptor computation modules and the top network as a similarity function, as described in Fig. 1 (a).

Basically **pseudo-siamese** networks have the same structure as the aforementioned siamese networks except that the weights of the two branches are not shared. This increases the num- ber ob adjustable parameters what results in more flexibility.

Ummon

Ummon is a Python library developed at the Institute for Optical Systems Konstanz to build and train neural networks. Since ummon was written in Theano[3] to some extent, computationally expensive arithmetic operations, such as convolutions, can be performed on a GPU instead of a CPU.

[3] http://deeplearning.net/software/theano/index.html

Experiments

At first we implemented the pseudo-Siamese network from [1] in Lasagne[4] to be able to draw a comparison. Afterwards we extended ummon so that we could implement the same network in ummon. The structure of both networks is illustrated in Fig. 1 (b).

To evaluate our networks we used a benchmark dataset from [4] which consists of three subsets - Liberty, Notre Dame and Yosemite - all of which contain more than 450,000 image patches (64 x 64 pixels). For our experiments we used 200,000 patches from each set.

To evaluate the prediction of out networks we used the binary hinge loss as cost function

$$L = \frac{1}{N} \sum_{i}^{N} \max(0, 1 - t_i p_i) \qquad (1)$$

where p_i indicates the output of the network and t_i 1 ; 1 indicates the corresponding la- bel. The standard SGD with a constant learning rate 0.1 was used for optimization. We used mini-batches with a size of 128 samples per batch. All datasets were trained for 600 epochs, which took about 5 hours per dataset.

Both trainings (Lasagne and ummon) were performed on a NVIDIA Tesla P100 16GB GPU.

Results

The evaluation of our experiments yields following results. As shown in Fig. 2 both networks, ummon and Lasagne, achieve approximately the same results. The error on test data ranges from 6%, when trained on Notre Dame data and evaluated on them, to 13%, when trained on Liberty data and evaluated on Yosemite.

Conclusion

The results of our experiments show that it is possible to learn a similarity function for comparing image patches directly from image data. Further experiments showed that performance can be increased up to an error rate of 4.53data if the entire data sets are being used for train- ing. An even better result could be expected by using an even larger data set with beyond one million image patches.

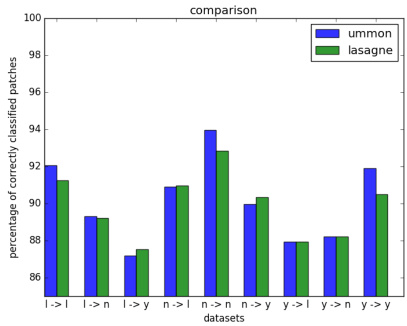


Abbildung 15: *Ummon* versus Lasagne on the aforementioned three datasets. The letter preceding the ar- row (-¿) indicates the dataset the neural network has been trained on. The letter succeeding the ar- row indicates the dataset to be predicted.

Literaturverzeichnis

[1] S. Zagoruyko and N. Komodakis. Learning to Compare Image Patches via Convolutional Neural Networks. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 2015

[2] B. Kugler. Lernen von Tiefeninformationen aus Stereobildern mit neuronalen Netzen. August 2017

[3] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. Interna- tional Journal of Computer Vision, 60:91â€“ 110, 2004

[4] M. Brown, G. Hua, and S. Winder. Discriminative learning of local image descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010

[4] https://github.com/Lasagne/Lasagne

Multifunktional-skalierbare generische InlineInspektion für flexible Fertigungsprozesse in vernetzten Produktionsanlagen

Matthias O. Franz, Georg Umlauf, Michael Grunwald und Matthias Hermann

Individualisierte, auf personifizierte Kundenbedürfnisse ausgerichtete Produkte sind in den kommenden Jahren eine DER technologischen Herausforderungen für die deutsche Industrie. Bei der Individualisierung klassischer, vorwiegend von Halbzeugen ausgehender Produktion stößt man aber an Grenzen. Daraus resultieren neue Herausforderungen für die fertigungsbegleitende (inline) Qualitätssicherung.

BMBF-Projekt MultiFlexInspect

Bearbeitung: *Prof. Dr. Matthias O. Franz*
Prof. Dr. Georg Umlauf

Laufzeit: *3 Jahre*

Zwei enthaltene Promotionsstellen:
1. Matthias Hermann, M.Sc.
2. N.N. <ausgeschrieben>

Abbildung 16: Überblick MultiFlexInspect.

Optische Sensorik für die flexible vernetzte Produktion

Eine leistungsfähige und starke Industrie ist in Deutschland die Basis für Wachstum, Wohlstand und qualifizierte Arbeitsplätze. Die hohe Dynamik der globalisierten Märkte und die immer kürzeren Innovationszyklen stellen jedoch auch etablierte und über lange Jahre erfolgreiche Unternehmen permanent vor neue Herausforderungen. Zukünftige Produktionssysteme müssen flexibel und adaptiv sein. Immer häufiger werden sie auch autonom agieren müssen. Damit einher geht ein immer größerer Bedarf an Informationen, auf deren Basis Maschinen ihr Umfeld und die zu bearbeitenden Objekte erkennen können. Die berührungslos arbeitenden Lösungsansätze der Photonik eignen sich in besonderer Weise zur flexiblen und schnellen Erfassung von Informationen über komplexe Zustände und Umgebungen. Das Potenzial der photonischen Sensorik - aufsetzend auf dem Stand der Technik - für den Einsatz in flexiblen und wandlungsfähigen Produktions- umgebungen mit teilweise autonom agierenden Maschinen zu erschließen, ist das Ziel dieser Fördermaßnahme. Gleichzeitig soll auch die visuelle Bereitstellung von Informationen für eine intuitive Anreicherung der Umgebungswahrnehmung im industriellen Umfeld mit zusätzlichen Informationen weiter vorangetrieben werden. In der flexiblen und vernetzten Produktion fällt der Informationsverarbeitung eine wesentliche Bedeutung zu. Entsprechende Kooperationen zur ganzheitlichen Betrachtung des Systems aus optischem Sensor und der zugehörigen Datenverarbeitung sollen unterstützt und weiter ausgebaut werden. Für die Forschungsarbeiten in 13 Verbundprojekten stellt das BMBF ca. 24 Millionen Euro zur Verfügung. Das Teilprojekt MultiFlexInspect (Förderkennzeichen: 13N14540) ist in Abbildung 1 dargestellt.

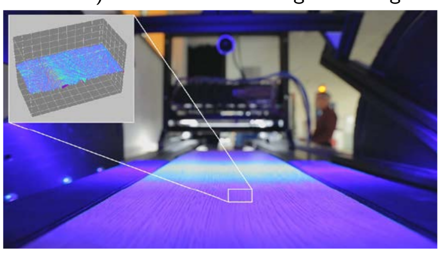


Abbildung 17: Inline-Analyse einer farbigen, texturierten 3D-Oberflächen (Quelle: Baumer Inspection GmbH).

Wandlungsfähigkeit in fertigungsbegleitender Qualitätssicherung ermöglichen

Individualisierte, auf personifizierte Kundenbedürfnisse ausgerichtete Produkte sind in den kommenden Jahren eine DER technologischen Herausforderungen für die deutsche Industrie. Neben der Automobilbranche zählt hier der industrielle Möbel- und Einrichtungsbau in Deutschland zu den Vorreitern. Bei der Individualisierung klassischer, vorwiegend von Halbzeugen ausgehender Produktion stößt man aber an Grenzen. Individualität wird oft 'erkauft' durch große Materiallagerflächen (oder Lager auf der Straße'), umfangreiche Werkzeugvarianten, erhebliche Bevorratung und letztlich immense Kapitalbindung und aufwendigste Logistik (OEM und Zulieferer). Hier können Verfahren zur schnellen und kostengünstigen Fertigung (Additive Manufacturing) dazu beitragen, dass das Individualisierungsniveau von Produkten weiter gesteigert wird und gleichzeitig ökonomische und ökologische Vorteile erreicht werden. Daraus resultieren neue Herausforderungen für die fertigungsbegleitende (inline) Qualitätssicherung. Vor diesem Hintergrund hat sich das Projekt MultiFlexInspect zum Ziel gesetzt, eine multi-funktionale Inspektionstechnologie zu erforschen und systemtechnisch zu demonstrieren, mit der 3D-Form und Maße sowie Farben und Oberflächenbeschaffenheiten eines Produktes in einem Arbeitsgang, inline und kosteneffizient bewertet werden können.

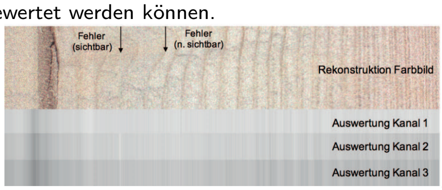


Abbildung 18: Beispielhafte Darstellung von wahrnehmungsorientierter Digitaldruckinspektion (vgl. [1]).

Durch menschliche Wahrnehmung inspirierte ganzheitliche Inspektion

Ausgangspunkt und technologische Basis des Projekts ist das System 3DPIXA aus dem Haus Chromasens - ein Zeilensensor für 3D Stereomesstechnik. Die 3D Messtechnik soll im Rahmen des Projektes mit einer präzisen geräteunabhängigen Farbmessung verknüpft werden, um mit derselben Sensorzeile parallel eine Erfassung von Spektralfarb-Informationen des Prüflings sowie Geometrie- und Maßdaten und sogar - basiert auf einer entsprechenden Auswertung - Oberflächenbeschaffenheit ermitteln zu können. Damit dies problemlos in dezentral organisierten und flexibel umkonfigurierbaren Anlagen möglich ist, sollen gleichzeitig Lösungen erarbeitet werden, die das System sehr adaptierfähig an verschiedene Applikationsspezifikationen macht. Einerseits sollen dazu automatisierte, eigenständig ausgeführte Kalibriermethoden erarbeitet werden. Ein weiterer Kern des dezentralen multifunktionalen Einsatzes soll die Implementierung einer künstlichen Intelligenz sein, um damit eine Skalierung und Adaptierung an wandlungsfähige Produktionsprozesse zu ermöglichen. Außerdem sollen durch Fa. Silicon Software entsprechend trainierte Algorithmen des Machine-Learning auf einen in den Sensor zu integrierenden, vor Ort programmierbaren, integrierten Schaltkreis (Field Programmable Gate Array; FPGA) implementiert werden, wo sie unter Echtzeit eingehende Bilder bearbeiten. Aufbauend auf einer solchen Inline-tauglichen Kontrolle von additiv gefertigten Produkten sollen bei Fa. Baumer zusammen mit den assoziierten Partnern Nobilia und Classen dynamische Prozessregelschleifen erarbeitet werden, welche direkt auf die Ergebnisse der intelligenten Inspektion aufbauen (vgl. Abbildung 2). Erstmalig sollen dabei bei der Bewertung nicht nur rein technische Aspekte, sondern auch von außen kommende Informationen (z. B. wahrnehmungsmotivierte Qualitätsbewertungen eines Menschen) in das Modell mit eingebunden werden, wofür die HTWG Konstanz verantwortlich zeichnet. Damit soll - einzigartig in der Inline-Inspektionstechnik - dem Aspekt Rechnung getragen werden, dass bei der Wahrnehmung von Defekten oft enorme Diskrepanzen in der Bewertung durch Mensch und Maschine bestehen, insbesondere für farbige texturierte 3D-Oberflächen. Heute kommt es nicht selten vor, dass Defekte auf Basis des menschlichen Eindrucks durchaus toleriert würden, während klassische maschinelle Inspektionsmodelle zur Fehlermeldung führen und somit zur Überklassifikation (vgl. Abbildung 3). Die vor Ort im Sensor vorhandene Intelligenz und insgesamt der multifunktionelle sensorische Ansatz sollen dazu beitragen, dass das neue System zu intelligenten, an der menschlichen Wahrnehmung orientierten Inspektionen befähigt wird.

Literaturverzeichnis

[1] M. Grunwald und Matthias O. Franz. Wahrnehmungsorientierte optische Inspektion von texturierten Oberflachen. In Heinrich C. Mayr, Martin Pinzger (Hrsg.): INFORMATIK, 2016.

Prediction of solar irradiation with Deep learning based on solar energy and meteorological time series

Robin Mattes, Martin Schall and Matthias O. Franz

Renewable energies are getting more important. But in case of solar energy it's difficult to predict future energy productions. In this work we show how a time series of sun energy observations and Deep Learning can be used to make predictions. With standard feed forward networks and in further steps with recurrent neuronal networks. [1]

Introduction

Renewable energy sources often depend on weather conditions such as wind and sun. Therefore they can't produce energy on a constant bases. That's a problem because it's difficult to know when enough energy is available from renewable sources. Especially when you want to trade in energy. This is not a problem for large operators as weather forecasts can be calculated with high precision and purchased. For small operators this is not economical because the profits are low and therefore a precise prediction is not worthwhile. In this work we focus specifically on small solar systems.The aim is to provide a low-cost prediction of solar radiation over 24 hours.In order to achieve this, publicly available data of solar radiation as well as weather conditions will be used. To calculate predictions from the existing data deep learning methods were used. In order to calculate forecasting from the available data the work from [1] was used as a base. It is shown that a prediction with a feedforward network is possible. In this solution, the time series data are selected using the embedded phase-space (EPS) method. This prevents the input data from being too similar. The method is used to create the training data for a dataset. The corresponding test data set consists of the following 24 data points. Thus, the network learns to make a prediction from the previous days (EPS) about the next 24 hours. The prediction is made with a feedforward network consisting of a hidden layer with 20 neurons and an output layer with 24 neurons. The size of the input layer depends on the length of the EPS data. National Renewable Energy Laboratory (NREL) records are used as a data basis. More exactly the data of the Weather Station New York Central PRK (ID725033). The time series include global solar irradiation (GHI). The years 2004-2005 are training data and 2006 is validation data. The average error of the prediction is calculated with root mean square error (RMSE) calculated over all 24 hours predictions.The result is shown in Figure 19. There is no significant difference between the different input values in the EPS model. Further we can see that the forecast is better in the first 5 hours. The ClearSky [2] curve serves as a comparison value because it is quite easy to calculate.

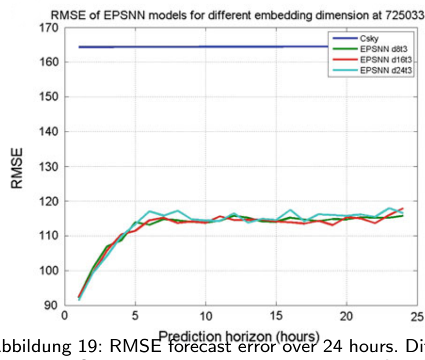


Abbildung 19: RMSE forecast error over 24 hours. Different EPS variables have similar results. The ClearSky straight line is plotted for comparison as it can be easily calculated. [1]

In the next step, we investigate how the approach could be improved. Recurrent neural networks (RNN) and LSTM networks were examined.

Methods and findings

It is investigated how the presented approach can be improved. For this purpose, further EPS variables were tested. The structure of the network was also investigated, including hidden layer sizes, gradient descent methods and activation functions. The result is that the EPS method is not needed for prediction. All improvements together have improved the RMSE by 4%. Figure 20 shows in green the prediction and in red the measured value over a period of 3 months. It can be seen that the prediction follows the measured values only slowly. In sections with low readings, the prediction takes a few days to align itself with the real values. Just like that with high values. Overall, the forecast curve is less extreme than the measured values.

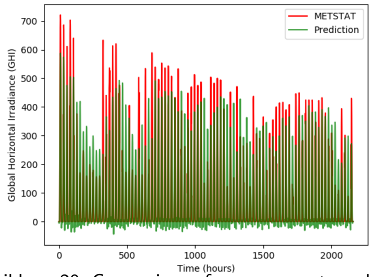


Abbildung 20: Comparison of measurements and predictions with the optimized feedforward network. In the time frame from 1 October 2005 to 31 December 2005 in hours.

RNN & LSTM

In the next step, it is investigated how to improve the prediction with recurrent neural networks. Classic RNNs as well as LSTMs (Long short-term memory) are looked at.
They are able to process not only the currently given input, but also inputs from the past due to their internal status. For this reason, it is investigated whether the average GHI error can be further improved with these networks. Furthermore, it is checked whether LSTMs are more suitable for prediction because they are able to detect long-term dependencies in the data. For this purpose, the findings on network structure and parameters of the RNNs are taken over and compared with their results. It is further investigated how many days are needed for a good forecast. 3 and approx. 180 (approx. half a year) days are proven to be the best timespan. Since weather is very random, most of the information for a forecast is probably from the first 3 days. At half a year, weather phenomena or seasons may also be taken into account in the forecast, thus improving the forecast.

Weather parameters

A qualitative enhancement of the model presented in [1] is the addition of weather parameters. Considered weather parameters are cloudiness, air pressure, humidity, temperature, wind direction and wind speed. In the first step, a training is carried out which includes all weather parameters as input data and the GHI time series of the sun as target data. This allowed us to show how much the parameters can tell us about solar energy. In the next step, only one weather parameter is added to the GHI time series and a forecast is created. Thus, it is possible to show which weather phenomena can contribute how much to the forecast. It turned out that the cloudiness has the greatest influence, which seems logical. Air pressure and humidity also have a significant influence. So far, only historical weather values have been used to calculate a forecast. The next step is to check how far the forecast of the weather data improves when a weather forecast is inserted into the training data. The existing weather data is used for this purpose. All the changes in the network and input data have led to a significant improvement in the prediction of the sun time series as shown in Figure 21.In comparison with Figure 20 it is clearly visible that the prediction (green) is much closer to the measured values (red) and is therefore more accurate.

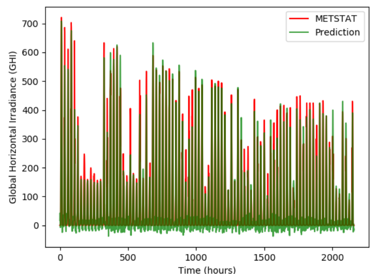


Abbildung 21: Comparison between prediction and measured values. The RMSE has improved by about 54% compared to the first feedforward network.

Next steps

In addition, it could be investigated whether the prediction can be further improved if a training sequence of several years is used. This would allow a network to identify more seasonal effects, if necessary, and thus to allow more accurate predictions. Other weather parameters such as precipitation probability and precipitation quantity could also contain further information.

Literaturverzeichnis

[1] Luigi Fortuna, Giuseppe Nunnari, and Silvia Nunnari. Nonlinear modeling of solar radiation and wind speed time series. *SpringerBriefs in energy* (, 2016.

[2] Pierre Ineichen and Richard Perez. A new airmass independent formulation for the linke turbidity coefficient. *Solar Energy*, 73(3):151–157, 2002.