Technische Universität München
TUM School of Engineering and Design

Automating the Programming of Robot-Based Optical 3D Measuring Systems

Alejandro Erick Magaña Flores

Vollständiger Abdruck der von der TUM School of Engineering and Design der Technischen Universität München zur Erlangung eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Tim C. Lüth

Prüfende der Dissertation:

1. Prof. Dr.-Ing. Gunther Reinhart

2. Prof. Dr.-Ing. Christoph Holst

Die Dissertation wurde am 09.04.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Engineering and Design am 21.11.2024 angenommen.

"The three great essentials to achieve anything worthwhile are, first, hard work; second, stick-to-itiveness; third, common sense."

Thomas A. Edison

Editors' Preface

In times of global challenges, such as climate change, the transformation of mobility, and an ongoing demographic change, production engineering is crucial for the sustainable advancement of our industrial society. The impact of manufacturing companies on the environment and society is highly dependent on the equipment and resources employed, the production processes applied, and the established manufacturing organization. The company's full potential for corporate success can only be taken advantage of by optimizing the interaction between humans, operational structures, and technologies. The greatest attention must be paid to becoming as resource-saving, efficient, and resilient as possible to operate flexibly in the volatile production environment.

Remaining competitive while balancing the varying and often conflicting priorities of sustainability, complexity, cost, time, and quality requires constant thought, adaptation, and the development of new manufacturing structures. Thus, there is an essential need to reduce the complexity of products, manufacturing processes, and systems. Yet, at the same time, it is also vital to gain a better understanding and command of these aspects.

The research activities at the Institute for Machine Tools and Industrial Management (*iwb*) aim to continuously improve product development and manufacturing planning systems, manufacturing processes, and production facilities. A company's organizational, manufacturing, and work structures, as well as the underlying systems for order processing, are developed under strict consideration of employee-related requirements and sustainability issues. However, the use of computer-aided and artificial intelligence-based methods and the necessary increasing degree of automation must not lead to inflexible and rigid work organization structures. Thus, questions concerning the optimal integration of ecological and social aspects in all planning and development processes are of utmost importance.

The volumes published in this book series reflect and report the results from the research conducted at *iwb*. Research areas covered span from the design and development of manufacturing systems to the application of technologies in manufacturing and assembly. The management and operation of manufacturing systems, quality assurance, availability, and autonomy are overarching topics affecting all areas of our research. In this series, the latest results and insights from our application-oriented research are published, and it is intended to improve knowledge transfer between academia and a wide industrial sector.

Rüdiger Daub Gunther Reinhart Michael Zäh

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Werkzeugmaschinen und Betriebswissenschaften (*iwb*) der Technischen Universität München.

Mein erster Dank gilt Herrn Prof. Dr.-Ing. Gunther Reinhart und Herrn Prof. Dr.-Ing. Michael F. Zäh für die fördernde Unterstützung meiner Arbeit und meiner persönlichen Entwicklung. Besonders bei Herrn Prof. Dr.-Ing. Gunther Reinhart möchte ich mich für die kontinuierliche Betreuung meiner Arbeit und das entgegengebrachte Vertrauen bedanken.

Des Weiteren möchte ich Prof. Dr.-Ing. Christoph Holst für das Koreferat und Prof. Dr. Tim Lüth für die Übernahme des Vorsitzes der Prüfungskommission danken. Zuletzt möchte ich Prof. Han Haitjema von der KU Leuven erwähnen, der mir die Möglichkeit gegeben hat, an seinem Institut im Rahmen meines Dissertationsvorhabens zu forschen.

Ich danke allen Kolleginnen und Kollegen des *iwb* für die wertschätzende Arbeitsatmosphäre, insbesondere der Themengruppe für Montagetechnik und Robotik für die lehrreiche Zeit. Hervorheben möchte ich hier Philipp Bauer, Thomas Rauh, Paul Geng, Jonas Dirr, Clemens Gonnermann und Daniel Gebauer, die stets offen für spannenden und konstruktiven Diskussionen rund um meine Dissertation waren.

Darüber hinaus möchte ich mich bei allen Studentinnen und Studenten für die großartigen Ideen und das Engagement bedanken, die mich im Laufe meiner Promotion begleitet haben. Lukas Schneider, Silvia Gebel, Thomas Altmann und Hang Wu verdienen an dieser Stelle eine besondere Erwähnung.

Asimismo, deseo expresar mi más sincero agradecimiento a mi familia y amigos, especialmente a mis padres y a mi hermana, quienes me apoyaron y motivaron incondicionalmente desde la distancia durante todo el proceso. Gracias.

Abschließend möchte ich einen besonderen Dank an meine Frau Elena richten, die mich all die Jahre unermüdlich unterstützt hat und zur Seite stand. Danke, dass du mich aus dem Alltag herausgeholt und mir neue Energie für den Forschungsalltag sowie zur Fertigstellung der Dissertation gegeben hast. Nun freue ich mich auf die Zeit mit unserer Tochter Valentina, die mir in der Abschlussphase meiner Promotion Kraft und Freude gegeben hat.

München, den 26.01.2025 Alejandro Magaña

Contents

Symbols

Economic Analysis

Symbol	Unit	Description
$c_{dep,a}$	€/a	linear depreciation per year
c_i	€	total investment costs
$c_{i,data}$	€	initial dataset generation costs
$c_{i,dev}$	€	solution modules software development costs
$c_{i,other}$	€	other investment costs
$c_{op,a}$	€/a	operator programming costs per year
$c_{op,h}$	€/h	operator programming costs per hour
$c_{other,a}$	€/a	other operative costs per year
c_{ref}^{f}	€/f	programming costs per inspection feature for reference use-case
$c_{sav,a}$	€/a	programming saving costs per year
$c_{semi-aut}^{f}$	€/f	programming costs per inspection feature for reference use-case with semi-automation
n_{break}	–	break-even number of inspecting features
P	a	payback period in years
$p_{ref,a}$	h/a	programming hours per year
$p_{sav,a}$	h/a	programming saving hours per year
p_{saving}	–	relative programming saving effort to reference use-case
T_{align}	min	time effort for automated alignment
T_{aut}	min	total programming effort based on full automated process
T_{par}	min	time effort for automated sensor parameterization
T_{ref}	min	reference total programming effort
$T_{semi-aut}$	min	total programming effort based on semi-automated process
T_{vp}	min	time effort for automated viewpoint planning

Machine Learning

Symbol	Unit	Description
$f(\boldsymbol{x})$	–	unknown function with a Gaussian distribution
g	–	activation function
$\mathcal{GP}$	–	Gaussian process
$k(\boldsymbol{x}, \boldsymbol{x}')$	–	kernel function between two input vectors
$m(\boldsymbol{x})$	–	mean function
R^2	–	coefficient of determination
RMSE	–	root-mean-square error
w	–	weight factor

(Symbols continued)

$\boldsymbol{x}$	–	input vector
x	–	input value
y	–	target value
$\boldsymbol{\mu}$	–	mean of a Gaussian distributon

Kinematic Model

Symbol	Unit	Description
C_0	–	World's frame coordinate system
C_{j_6}	–	robot's sixth joint frame coordinate system
C_r	–	Robot's frame coordinate system
C_s	–	Sensor's lens frame coordinate system
C_{tcp}	–	Sensor's TCP frame coordinate system
C_w	–	Workpiece's frame coordinate system
$\boldsymbol{p}_0^w$	–	workpiece's pose vector in world coordinate system
$\boldsymbol{p}_r^{j_6}$	–	six joint pose vector in robot's coordinate system
$\boldsymbol{p}_r^{tcp}$	–	sensor's TCP pose vector in robot's coordinate system
p_{tcp}^f	rad	feature's pitch rotation in the TCP coordinate system
r_{tcp}^f	rad	feature's roll rotation in the TCP coordinate system
$\mathbf{T}_0^r$	–	homogeneous transformation matrix between robot's and world coordinate systems
$\mathbf{T}_0^{tcp}$	–	homogeneous transformation matrix between sensor's TCP and world coordinate systems
$\mathbf{T}_0^w$	–	homogeneous transformation matrix between workpiece and world coordinate systems
$\mathbf{T}_{j_6}^{tcp}$	–	homogeneous transformation matrix between sensor's tcp and sixth joint axis coordinate systems
$\mathbf{T}_r^{j_6}$	–	homogeneous transformation matrix between sixth joint axis and robot coordinate systems
$\mathbf{T}_w^{tcp}$	–	homogeneous transformation matrix between workpiece and sensor's TCP coordinate systems
w_{tcp}^f	rad	feature's yaw rotation in the TCP coordinate system
x_{tcp}^f	mm	feature's x coordinate in the TCP coordinate system
y_{tcp}^f	mm	feature's y coordinate in the TCP coordinate system
z_{tcp}^f	mm	feature's z coordinate in the TCP coordinate system

Sensor Model

Symbol	Unit	Description
d_s	mm	sensor's working distance to middle plane of measurement volume
h_s^{far}	mm	far plane of sensor's measurement volume

(Symbols continued)

h_s^{middle}	mm	middle plane of sensor's measurement volume
h_s^{near}	mm	near plane of sensor's measurement volume
i_{avg}	–	average image intensity around a feature in grayscale image, $i_{avg} \in [0-255]$
$I_s^{i,j}$	–	image intensity at pixel x_i, y_j
L_a	–	direct ambient light intensity
L_o	–	ambient light intensity reflected by an object
L_p	–	projector light intensity reflected by an object
p_{norm}	–	normalized number of surface points
t_{exp}	ms	exposure time of optical sensor
z_s	mm	distance between sensor image plane and surface
γ_s	–	camera sensitivity
θ_s^x	rad	width field of view angle
ρ_s	–	surface reflectivity of an object
σ_s	–	sensor noise
ψ_s^y	rad	length field of view angle
ϕ_s	rad	sensor incidence angle at a surface

Viewpoint Model

Symbol	Unit	Description
$\widetilde{C}$	–	viewpoint constraint
F	–	set of inspection features $f \in F$
f	–	inspection feature
G	–	subset of inspection features $G \subseteq F$
v	–	viewpoint

Acronyms

Short	Description
AI	artificial intelligence
ANN	artificial neural network
CAD	computer-aided design
CMM	coordinate measuring machine
CNN	convolutional neural network
CSG	constructive solid geometry
$\mathcal{C}$-space	feature-based constrained space
DH	*Denavit-Hartenberg*
DNN	deep neural network
DoF	degree of freedom
DRM	Design Research Methodology
FOV	field of view
${}^{G}\mathcal{C}$-space	feature cluster constrained space
GP	Gaussian process
HDR	high dynamic range
iwb	*Institute for Machine Tools and Industrial Management of TU Munich (German: Institut für Werkzeugmaschinen und Betriebswissenschaften der TU München)*
LoA	level of automation
LoA_i	cognitive/information level of automation
LoA_m	mechanical/physical level of automation
ML	machine learning
PUB	publication
ROMS	robot-based optical 3D measuring system
RQ	research question
SCP	Set Covering Problem
SE	squared exponential kernel
SLS	structured light sensor
SM	solution module
SME	small and medium-sized enterprise

(Acronmys continued)

TCP	tool center point
TUM	Technical University of Munich
VGP	Viewpoint Generation Problem
VPP	Viewpoint Planning Problem

Chapter 1

Introduction

1.1 Motivation

Globalization, demographic change, emerging technologies, dynamic product life cycles, and knowledge management are some megatrends that have influenced the producing industry over the last decade and will continue defining its future development (ABELE and REINHART 2011, p. 10). To overcome these challenges, ABELE and REINHART (2011, p. 34) identify four target variables that producing companies must address to remain competitive: *manufacturing costs, speed* (*innovation, process,* and *product launch*)*, changeability*, and *quality*. For the field of quality assurance, SCHMITT et al. (2011) identified furhter challenges in production metrology[1], such as reducing measurement times, enhancing measurement accuracy, facilitating faster commissioning and integration, increasing adaptability to different measurement tasks, and developing holistic systems that encompass multiple steps in metrology processes (e.g., planning, acquisition, processing, and evaluation).

In particular, optical three-dimensional (3D) measurement technologies have positioned themselves as an attractive asset to address some of the challenges mentioned above and as a cost-effective driving force to accelerate the digitization of manufacturing (CATALUCCI et al. 2022; GAO et al. 2015; SCHMITT et al. 2016). Using optical triangulation techniques, 3D optical sensors can scan the surface of an object of interest without contact and provide an accurate discretized 3D representation of it, known as *point clouds*. These 3D measurements can then further used as input data to perform and support different tasks, e.g., measurement, monitoring, handling, and navigation (CHEN et al. 2011; PEUZIN-JUBERT et al. 2021).

Optical 3D sensors have demonstrated their significance as a key resource for various metrological tasks and within digital production. However, their effectiveness in more complex production tasks such as large-scale metrology applications is limited when they are fixed in position. For example, the inspection of common

[1] According to PFEIFER (2002, p.15), production metrology encompasses all measurement and testing activities in connection with the industrial development of a process.

components in the automotive, energy and aerospace sector (e.g., car chassis, wind turbines, or fuselage of airplanes), require multiple acquisition views that demand a positional flexibility of up to six *degrees of freedom* (DoFs) of the measurements systems (SCHMITT et al. 2016). The demand for more flexible systems that permit the free positioning of sensors and automation of metrology tasks to acquire large-scale objects has driven the industry towards the use of *robot-based optical 3D measuring systems* (ROMSs). Such systems utilize a programmable industrial robot to position a 3D sensor. Figure 1.1 depicts a simplified representation of a ROMS and its core components.

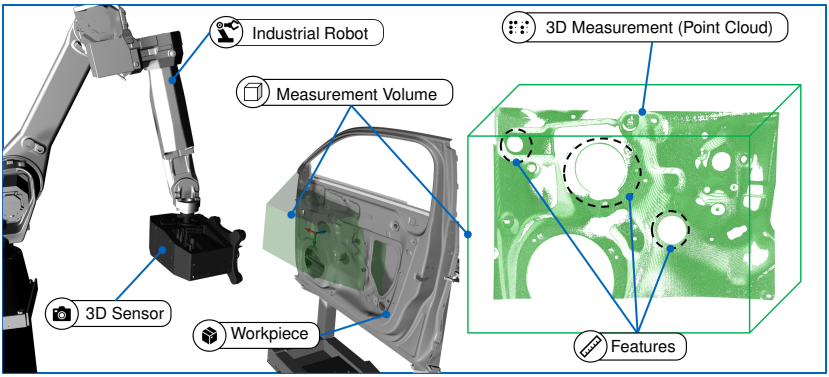


Figure 1.1: Simplified representation of a *robot-based optical 3D measuring system* (ROMS) and its fundamental components utilized for automating large-scale metrology tasks. The 3D sensor, an optical measuring device, is positioned by the robot to acquire multiple 3D measurements (point clouds) of the workpiece's surface. The point clouds are used to evaluate the quality of a workpiece by analyzing the geometric characteristics of features.

The inherent flexibility and automation capabilities offered by robotic manufacturing systems have proven effective in addressing some of the pressing issues that manufacturing industries face, such as quality, productivity, and competitiveness (MÜLLER and KUTZBACH 2019, p. 534). However, their programming is still considered a resource-intensive and complex activity that counteracts the benefits of robot-based production systems, especially for producing small batch sizes or products with a high variability (DIETZ 2019). This characteristic contradicts the requirement of manufacturing systems demanding "zero setup times" and minimal human intervention to be cost-effective (ABELE and REINHART 2011, p. 34; PFEIFER and SCHMITT 2010, p. 314). According to a report by the THE BOSTON CONSULTING GROUP (2015), the system engineering costs allocated for commissioning and programming robot-based systems correspond as much as one-third of the total investment costs. Especially programming ROMS to automate metrology tasks is considered a challenging process due to the increased system complexity and required multi-domain expertise in robotics, machine vision, optics, and metrology (CATALUCCI et al. 2022; BATCHELOR 2012, p. 16)

The resource-intensive programming of ROMS and the the scarcity of skilled professionals in high-wage countries are considered limiting factors for their cost-effective deployment (LITZENBERGER and KUTZBACH 2016, p. 431). The productivity of production systems can only be increased if they are able to adapt to changes or new tasks with less effort (HAMMERSTINGL 2020, p. 169; ULRICH 2018, p. 1).

The present thesis addresses these demands by outlining three solution modules that combine fundamental techniques from the classical robot domain, computer vision, and machine learning methods to improve the programming of ROMS. The outcomes of this thesis aim to decrease the programming effort required to support ROMS. This will benefit two groups: firstly, SMEs that need adaptable production systems with swift ramp-up times (DIETZ 2019, p. 1); and secondly, companies heavily reliant on robot-based solutions, such as automotive manufacturers. By reducing programming resources, these companies can enhance their productivity.

Moreover, this thesis strives to establish a formalized understanding of the knowledge necessary for automating programming of ROMSs. By identifying knowledge gaps, customized solutions can be devised with consideration of system constraints. It is crucial for this knowledge to be modeled as generically as possible to facilitate transferability to comparable applications. Complete automation liberates operators from repetitive tasks, enabling them to channel their potential into more innovative endeavors.

1.2 Scope of the Work and Problem Statement

To define the scope of this research, first Subsection 1.2.2 introduces the core components of a ROMS. Then, an exemplary metrology task is introduced in Subsection 1.2.2, which will be used as reference throughout this thesis.

Furthermore, Subsection 1.2.3 presents an analysis to estimate the current *level of automation* (LoA) of ROMSs and identify the subtasks with the highest potential for improving the programming of ROMSs.

1.2.1 Robot-Based Optical 3D Measuring Systems

A ROMS is a measuring system consisting of an *industrial robot* that is used to position a *range sensor* for performing metrology tasks (see Figure 1.1). Hereby, the sensor acquires a point cloud that represents a portion of the surface of the *workpiece*. The point cloud is used to measure the individual *features* of the workpiece and assess its actual geometric properties (WECKENMANN 2012, p. 305). The following subsections offer a detailed description of the core components—robot and sensor.

1.2.1.1 Industrial Robot

The norm DIN EN ISO 10218-1 (2012) defines an industrial robot as an automatically controlled, multi-purpose, fixed or movable manipulator that is freely programmable in three or more axes and is used for automation. Each robot has a defined number of DoFs which indicates how many independent motions (up to three translational and three rotational) a robot can perform within the robot workspace. For example, in the context of ROMSs, six-axis serial industrial robots are often used to automate large-scale metrology tasks due to their high flexibility for positioning the sensor up to six DoF within large workspaces. Subsection A.1.1 provides a description of the definition of the robot's workspace and kinematic model.

1.2.1.2 Range Imaging Sensor

The DIN EN ISO 10360-13 (2023) defines a range imaging sensor as a non-contact surface measuring sensor which collects 3D spatial data from the surface of an object of interest. According to BESL (1988), a range imaging sensor represents an enclosed system that combines hardware and software modules capable of producing a range image of the surface of an object of interest. A range image refers to the generated point cloud which is the output of the sensor after triggering a measurement action.

A point cloud represents a collection of 3D points in the Euclidean space, where each point corresponds to a surface point of the measured part. The computation of depth information is acquired using different optical measuring principles, e.g., triangulation, light intensity, or time of flight (BEYERER et al. 2016, p. 229). Fig. 1.2 depicts a simplified taxonomy and a graphical representation of some acquisition principles.

A prominent example within the category of triangulation principles is active range imaging sensors, which consist of at least one camera and one projector. The projector displays a light pattern at the acquisition scene. Then, the 3D coordinates of the surface points can be calculated through triangulation in a reference coordinate system if the corresponding camera image pixels for the projecting pattern can be estimated (GÜHRING 2002, p.16). On the contrary, if a pixel cannot be uniquely identified due to improper image saturation, the corresponding depth information cannot be estimated (LI and KOFMAN 2014; LIN et al. 2017a). Subsection A.1.2.1 describes the technical rationale behind acquiring 3D data based on active sensing.

In the last two decades, active sensors have gained popularity within large-scale metrology applications due to their low investment cost, increasing accuracy, large measurement volume, and high measurement speeds (KEFERSTEIN and MARXER 2015, p. 234; ULRICH 2018, p. 7; SANSONI et al. 2009; GÜHRING 2002, p. 16). For these reasons, this thesis focuses on automating metrology tasks through the use of active sensors.

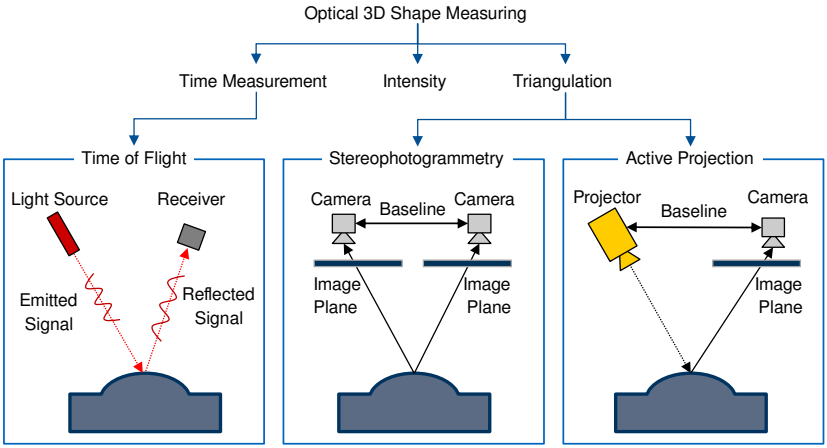


Figure 1.2: Overview of different acquisition principles for range imaging according to BEYERER et al. (2016, p. 229)

1.2.2 Metrology Task

The DIN EN ISO 9000 (2015) norm defines *quality* as the degree to which a set of different characteristics (e.g. physical, sensory, behavior) of a workpiece fulfills a set of defined requirements. Hence, one of the main tasks of dimensional metrology tasks is to verify the conformity of geometric *inspection features* to certain product specifications, such as size, form, location, and orientation (PFEIFER 2002, p. 254; HOCKEN 2012, p. 58; WECKENMANN 2012, p. 25). A feature is defined by its type, position, orientation, and geometric shape. Descriptors for topology properties include for example the radius of a hole or the lengths of a square. The ISO 1101 (2017) norm provides a detailed specification of different geometric characteristics.

Accurately estimating these features is considered a critical step in any manufacturing process necessary to evaluate tolerances within the product assembly and, ultimately, ensure product functionality (SCHMITT et al. 2016; PFEIFER and SCHMITT 2010, pp.1–2). Traditionally, *coordinate measuring machines* (CMMs) are used to determine these properties with a high accuracy based on tactile sampling of a finite set of surface points. However, the high setup effort, costs, and measurement times associated with CMMs made them cost-ineffective for the acquisition of large-scale workpieces with multiple inspection features (ULRICH 2018, pp. 2–3; PFEIFER and SCHMITT 2010, p. 315).

The technological advances of 3D optical sensing in terms of accuracy and processing speeds, combined with the flexibility of an industrial robot, make ROMSs a promising technology for automating these tasks. Automated acquisition of multiple point clouds enables the metrology evaluation of large-scale objects, as shown in the left side of Fig. 1.3. However, acquiring large surfaces requires

multiple viewpoints due to the sensor's field *field of view* (FOV) limitations (WECKENMANN 2012, pp. 17–18). In addition, setting the sensor parameters to acquire valid measurements is considered a resource-intensive process that requires skilled personnel (CATALUCCI et al. 2022). This makes the programming of ROMSs for feature-based dimensional metrology a challenging and time-exhaustive task with a high potential for automation.

For these reasons, the present thesis uses a reference dimensional metrology task to demonstrate the programming complexity of ROMSs and the need for its automation. The reference metrology task involves the inspection of a sheet metal car door with over 500 different inspection features (e.g., edges, pockets, holes, slots, and spheres). Fig. 1.1 visualizes on the right side some of these exemplary features. Car doors are well-known benchmarking probing objects for evaluating technologies used within large-scale metrological tasks due to their size, topological complexity, feature density, and variability (BAUER et al. 2021a; TEKOUO MOUTCHIHO 2012, p. 124).

The measurement task is completed when all features are acquired with a valid *point cloud quality*. In the context of this dissertation, point cloud quality refers to the local point cloud density around a feature. The point cloud quality is intended to be used as a metric to represent whether the acquired measurement can be used to evaluate an individual feature. The point cloud quality is not related to the accuracy of the sensor and can only be affected by the sensor parameters, position, or environmental conditions. For instance, the measurement depicted on the right side of the Fig. 1.3 suggests that while the quality of the point cloud surrounding certain features is acceptable for evaluation purposes, it is insufficient for others.

Moreover, it is assumed that the nominal position and orientation of all features, sensor imaging properties (e.g., measurement volume, resolution, working distance), the robot kinematic model are roughly known, and that the measurements are acquired statically.

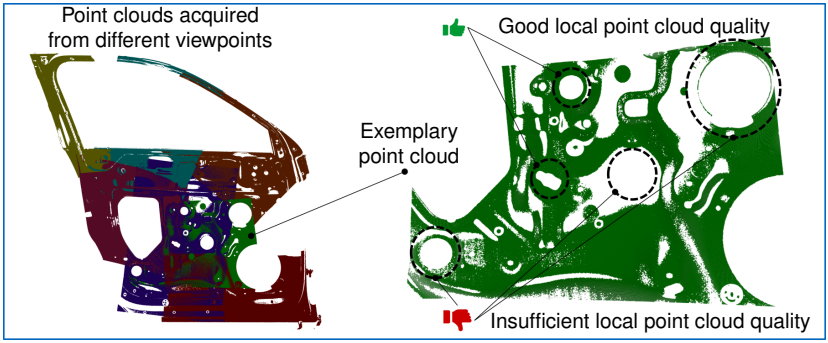


Figure 1.3: Left: Multiple point clouds acquired from different viewpoints. Right: An exemplary point cloud with diverse features, exhibiting different local point cloud qualities.

1.2.3 Level of Automation of ROMS

1.2.3.1 Classification

In the literature, a wide range of taxonomies have been proposed to classify the LoA of production systems (FROHM et al. 2008; GAMER et al. 2019; VAGIA et al. 2016). The present thesis follows the framework from FROHM (2008) to estimate the LoA of ROMSs. The study of FROHM et al. (2008) introduces two independent categories to estimate the *mechanical/physical level of automation* (LoA_m) of production systems, at a LoA_m and at a *cognitive/information level of automation* (LoA_i). The mechanical autonomy of a system to interact within its production environment is measured by the LoA_m, which includes the ability to handle a reference workpiece or position a required work resource. By contrast, LoA_i evaluates the amount of information the system requires from a user to perform a task. In addition, it also quantifies the cognitive capabilities of the system to control the target process or generate alarms. Within each category, FROHM et al. (2008) introduce seven levels of automation, ranging from a manual process (Level 0) to a fully automated process (Level 7), i.e. autonomous systems[2].

1.2.3.2 Subtasks with the Highest Automation Potential

According to GRANELL et al. (2007), the reference metrology task (see Subsection 1.2.2) is further analyzed and broken down into subtasks to identify those with the highest automation potential. For this purpose, the steps for performing a metrology task according to PFEIFER and SCHMITT (2010, pp. 13–16) are considered: inspection planning, data acquisition, data evaluation, process control, capability analysis, and monitoring of measurement devices. Among these, the following tasks were identified as relevant for improving the automation of ROMSs:

1. Inspection planning: This task considers the definition of the inspection plan (PFEIFER and SCHMITT 2010, pp. 14–15). According to the norm DIN 55350 (2021), an inspection plan integrates all necessary information to perform an inspection task, e.g., feature properties and characteristics, measuring tolerances, equipment, and measuring strategies. Hence, the inspection plan provides the basis of any inspection order and must be fully specified beforehand.

2. Data acquisition: Collecting measurement data based on the inspection plan is the central task of production metrology. To accomplish this task, it is necessary to consider the capabilities of the measurement equipment in terms of accuracy and speed, as well as the degree of automation. (PFEIFER and SCHMITT 2010, p. 15)

[2] WATSON and SCHEIDT (2005) define autonomous systems as agents that are able to adapt their behavior during operation to unforeseen events without manual intervention.

3. Data evaluation: This step is concerned with condensing the acquired data into static characteristic values (PFEIFER and SCHMITT 2010, p. 15). For example, the acquired data is used to evaluate the nominal's workpiece geometry using feature extraction algorithms to quantify the deviation from the product specifications (PFEIFER and SCHMITT 2010, pp. 196–197).

In the context of programming metrology tasks using ROMSs, the *viewpoint planning* as a part of the measurement strategies within inspection planning, and *parameterizing measurements* during data acquisition were identified as highly time-consuming subtasks that required active expert intervention. These subtasks were confirmed as time consuming in the study by BAUER et al. (2021a), which evaluated the economic potential of ROMSs. In addition, *workcell calibration,* was recognized as a generic, preparatory task in the context of robot programming that requires human intervention and therefore has potential for automation. Figure 1.4 provides a simplified graphical representation of the selected subtasks and their relationships. These are briefly described below.

1. Workcell calibration: This step addresses the estimation of the position and orientation of the inspecting workpiece or any other artifact, such as fixtures within the robot workspace that may collide with the robot or affect the measurement acquisition. Calibration of the workcell is a mandatory and primarily manual process which the operator must perform before executing a programmed routine (BRECHER and WECK 2021, p. 594).

2. Viewpoint planning: This subtask addresses the selection of a sufficient number of viewpoints (sensor poses) for capturing all features. The literature poses this task as a multi-objective optimization problem referred to as the *Viewpoint Planning Problem* (VPP) (PEUZIN-JUBERT et al. 2021). The goal for the operator or automation logic is not solely to reduce the number of viewpoints, but also to fulfill various constraints to ensure that all features are captured. Estimating valid viewpoints is a complex and challenging task that requires specialist expertise (CATALUCCI et al. 2022; GOSPODNETIĆ et al. 2022).

3. Measurement parameterization: The adequate parameterization of the sensor is a critical requirement to guarantee the 3D sampling of the object's surface when using optical imaging devices. In particular, the proper setting of the image exposure is a non-trivial task requiring the consideration of diverse influencing factors, e.g., camera parameters, surface material, the spatial relationship between camera and workpiece, and external lighting sources (EKSTRAND and ZHANG 2011; ZHANG 2020). For this reason, selecting a proper exposure time remains a challenging and time-consuming task performed by experts, which rely on their domain knowledge and spatial understanding to achieve proper adjustment.

Further relevant subtasks within metrology applications address the registration of point clouds and evaluation of the acquired data. However, efforts to address the challenges in these processes are focused on improving measurement uncertainty, increasing acquisition speeds, and effectively managing large data sets (BAUER et al. 2021b; CATALUCCI et al. 2022; ULRICH et al. 2015). For these reasons, these subtasks fall outside the scope of the present dissertation.

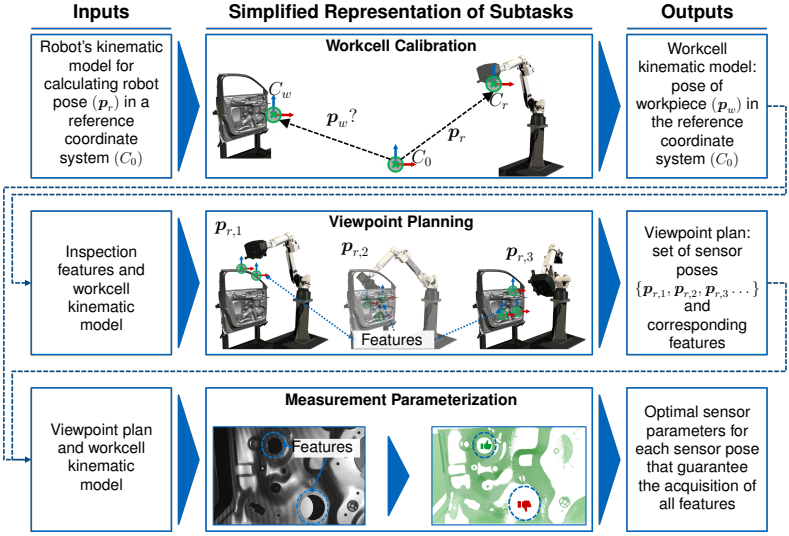


Figure 1.4: Overview of the identified subtasks with the highest potential for automating metrological tasks using ROMSs: workcell calibration, viewpoint planning, and measurement parameterization.

1.2.3.3 Current LoA of ROMS

FROHM et al. (2008) categorize industrial robot systems as flexible workstations that can be reconfigured for different tasks and perform a programmed routine without human intervention. For this reason, the authors assign such systems an LoA_m of six. This thesis also classifies the mechanical automation level of ROMSs at the same level assuming that the industrial robot can freely position the sensor without mechanical human assistance.

To operate a ROMSs at a LoA_m of six, it must be assumed that all necessary subtasks have been adequately programmed and parameterized in advance (cf. Subsection 1.2.3.2). Since the level of cognitive automation achieved by current solutions for the programming of metrology tasks still requires a considerable amount of manual intervention, this thesis assumes a teaching level of LoA_i of the order of three for current ROMSs.

1.3 Research Methodology

This section outlines the research methodology used in this thesis. Based on the findings of the previous section, this research presents first its overall goal and the necessary *research questions* (RQs) to achieve it.

1.3.1 Thesis Objective and Research Questions

Robot-based manufacturing systems and optical 3D sensors have been identified as key technologies to address some of the challenges facing producing companies, such as, production flexibility, adaptability, and labor shortages. However, the effort required for their commissioning and programming prevents their utilization and affects their economic benefits and productivity. For these reasons, this thesis aims to increase the level of automation for the programming of ROMSs.

This dissertation suggests that the LoA of ROMSs can be increased up to a *supervision* stage ($LoA_i = 5$) by automating the identified subtasks in Subsection 1.2.3.2. Within this level, the operator acquires a supervising role, where the system would only demands the user's attention in case of anomalies (FROHM et al. 2008). This enhancement offers the potential to elevate the level of automation in programming ROMSs, thereby augmenting their productivity and economic feasibility.

To tackle this objective, the current research focuses on the automation of three subtasks for facilitating the programming effort of metrological applications using ROMSs. To assess this goal and guide the present research, the following RQs are introduced:

RQ1 *Workcell Calibration: How can the workcell calibration be automated by using images captured by the optical sensor of ROMSs?*

RQ2 *Viewpoint Planning: How do viewpoint constraints affect sensor positioning, and how can the minimum number of viewpoints required to capture all features be estimated?*

RQ3 *Exposure Time Optimization: How do image exposure and the spatial correlations between features and the sensor affect the successful acquisition of 3D measurements, and how must these be considered for setting a proper exposure time?*

1.3.2 Research Design and Thesis Structure

In accordance with the classification schema for scientific disciplines as proposed by ULRICH and HILL (1976), the current study is situated within the domain of

the real sciences. This domain necessitates a distinction between fundamental research and applied research. The methodologies employed by RQ1 and RQ3 align with the characteristics of applied research. In contrast, RQ2 demands a comprehensive and fundamental exploration into the various constraints that influence sensor positioning. In order to adequately address the knowledge gaps related to all the research questions, it is necessary to take an interdisciplinary approach and leverage insights from various fields such as production engineering, metrology, robotics and computer vision. According to (KUBICEK 1976, p. 13), modeling complex systems requires an iterative heuristic methodology to gain a holistic understanding of systemic interactions and impacts. Accordingly, this thesis adheres to an iterative process of empirical observations based on from simulations, experiments, and data gathering to augment and validate the acquired knowledge.

In order to effectively acquire the knowledge necessary to address the research questions at hand, this thesis utilizes the *Design Research Methodology* (DRM) introduced by (BLESSING and CHAKRABARTI 2009) for structuring research projects in applied fields like engineering. The DRM provides the necessary framework for conducting empirical, iterative research. Therefore, the present research follows the *Development Support* research type of the DRM to define the methods employed within the following fundamental stages, i.e., *Research Classification, Descriptive Study I, Prescriptive Study, and Descriptive Study II* (BLESSING and CHAKRABARTI 2009, p. 61). The research structure suggested by the DRM provides the inherent outline of this thesis and is summarized in Tab. 1.1.

In the first stage, the *Research Classification* foresees the definition of research objectives and questions (see Subsection 1.3.1) based on the description and assumptions of an exemplary initial situation of the existing problem. The definition of a reference metrology task in Subsection 1.2 serves this purpose and enlightens the potential for improving the automation of ROMS. Moreover, this stage also considers the formulation of some criteria, such as the defined autonomy levels in Subsection 1.2.3 for evaluating the research outcome.

The next stage, *Descriptive Study I,* addresses the specification of the problem based on an exhaustive literature analysis. First, Chapter 2 presents related research addressing the automation of the identified subtasks. The literature review within this stage supports the present thesis's relevance and discusses the identified research gaps. Moreover, Section A.1 introduces the core components of ROMSs and the mathematical foundations from the machine learning domain.

The third stage, the *Prescriptive Study*, comprises the conceptualization and implementation of three *solution modules* (SMs) to enhance the programming of each identified subtask. Since the present dissertation follows a publication-based structure, the conceptualization and implementation of each SM are summarized in Chapter 3. A more exhaustive description of the concepts is given in the four publications PUB1–PUB4, which investigate the formulated RQs as follows:

1. RQ1 is addressed in PUB1: MAGAÑA *et al. (2020a), „PoseNetwork: Pipeline*

for the Automated Generation of Synthetic Training Data and CNN for Object Detection, Segmentation, and Orientation Estimation“.

2. RQ2 is addressed in
    - PUB2*:* MAGAÑA *et al. (2023a), „Viewpoint Generation Using Feature-Based Constrained Spaces for Robot Vision Systems“ and*
    - PUB3*:* MAGAÑA *et al. (2023b), „Viewpoint Planning for Range Sensors Using Feature Cluster Constrained Spaces for Robot Vision Systems“.*
3. RQ3 is addressed in PUB4*:* MAGAÑA *et al. (2023c), „Exposure Time and Point Cloud Quality Prediction for Active 3D Imaging Sensors using Gaussian Processes“*

In the last stage, the *Descriptive Study II* foresees a holistic evaluation of the technical and economic impact of the outlined solutions. Hence, Chapter 4 summarizes first the technical contributions for automating metrology tasks using ROMSs, as well as their effectiveness and computational efficiency to ease their programming effort. Moreover, the economic benefits and limitations of the proposed solutions are analyzed and compared to a manual programming of ROMSs. Finally, in the last step of this stage, a summary of the obtained results and the research outlook regarding the automation of ROMSs is provided in Chapter 5.

Table 1.1: Overview of the present thesis structure and deliverables aligned to the stages of the DRM (BLESSING and CHAKRABARTI 2009). The main contributions of the publications PUB1–PUB4 are placed within the prescriptive stage of the research methodology.

Research Stage	Deliverable
Research Classification (Review-Based)	- Identification of subtasks with the highest automation potential for ROMSs (Subsection 1.2) - Definition of research objectives (Subsection 1.3.1)
Descriptive Study (Review-Based)	- Problem understanding (Chapter 2) - Identification of research gap and subtasks challenges (Subsections 2.2.5, 2.3.5, and 2.4.3)
Prescriptive Study (Comprehensive)	- Conceptualization of solution modules (Sections 3.1.1, 3.1.2, and 3.1.3) - Implementation of solution modules (Subsections 3.2 (PUB1), 3.3 (PUB2), 3.4(PUB3), and 3.5(PUB4))
Descriptive Study II (Initial)	- Holistic technical and economic evaluation (Chapter 4) - Summary and critical review (Chapter 5)

Chapter 2

State of the Art

This section provides a summary of research activities that have presented solutions for automating or improving the identified subtasks during the programming of *robot-based optical 3D measuring systems* (ROMSs). First, to provide a comprehensive understanding of the programming methods used for RVS in manufacturing systems, Subsection 2.1 offers a general overview. The following Subsections 2.2–2.4 outline the efforts made to automate the individual subtasks selected for this thesis.

2.1 Robot Programming Methods

The goal of programming robot-based manufacturing systems is to generate a robot program that consists of a set of robot poses, collision-free trajectories between poses, and fine-tuned system parameters that ensure the automated execution of a targeted manufacturing process.

A variety of methods exists that combine different hardware and software solutions to generate robot programs. For most approaches, a distinction is made between online and offline programming (REINHART et al. 2018, pp. 149–150). This subsection provides an overview of the main characteristics of some common online and offline programming methods. A more comprehensive overview and classification of programming methods is given by BRECHER and WECK (2021, pp. 589–601) and VOGL (2009, pp. 12–26).

2.1.1 Online Programming

Online programming methods are performed on-site using the physical system under real process conditions. Hereby, an expert programs the required routine by positioning the end-effector in the required robot poses and trajectories using a teach-panel (BRECHER and WECK 2021, p. 590). To ensure process stability, the operator is also responsible for adjusting process or system parameters, e.g., camera parameters, welding parameters, and gripper commands.

2.1.2 Offline Programming

In contrast, offline programming methods do not require the use of the physical system to create a robot program. Instead, the robot poses and motions are written textually in a robot programming language (text-based) or generated with the help of a kinematic simulation. For instance, text-based programming is tightly coupled to the robot hardware, and the automation of process is modeled by a set of control commands (VOGL 2009, p. 21). On the contrary, simulation-based methods use a visualized kinematic model of the workcell to program motions that are translated into robot routines. The kinematic model of the workcell integrates the robot's kinematic model (see Subsection 1.2.1.1) and the spatial relationships of all relevant[1] physical system components within it (BRECHER and WECK 2021, p. 597).

Although simulation software tools have shown great potential for reducing programming effort, a high level of expertise is still required to generate and validate robot poses and trajectories (PAN et al. 2012; VOGL 2009, p. 22). For this reason, implicit or task-oriented programming approaches have emerged in the context of automated programming as an evolution of offline programming based on kinematic simulations. These methods aim to abstract and formalize the necessary domain knowledge (e.g., process, workpiece, robot, end-effector) to generate robot programs. The knowledge is formalized in the form of a logic, program, or algorithm that ultimately generates a suitable sequence of robot poses and trajectories. Thus, the robot programmer only needs to specify the task instead of programming robot joints or poses. (VOGL 2009, p. 24; BRECHER and WECK 2021, p. 600; TEKOUO MOUTCHIHO 2012, p. 28)

In recent years, task-oriented programming has gained popularity because it increases the flexibility and robustness of programming in manufacturing systems while reducing the effort and expertise required (BACKHAUS 2016; HAMMERSTINGL 2020; STENMARK and MALEC 2015; TEKOUO MOUTCHIHO 2012). However, the proper and exhaustive formalization of domain knowledge remains a challenging task when developing a programming framework for the automated generation of robot programs.

2.2 Workcell Calibration

Having an accurate kinematic model of the workcell is an essential requirement for offline programming approaches that aim to automate the generation of robot routines (see Subsection 2.1.2). Initially, a preliminary kinematic model is created from a rough workcell layout or by manually measuring the spatial relationships between relevant components. The following step involves a calibration process to determine any deviation from the initially estimated position

[1] In this context, a relevant component refers to any object that could impact the programmed motion considering the physical system or potentially collide with the robot during its path.

and orientation of workcell components like workpieces and collision objects (HÄGELE et al. 2016, p. 1398). First, Subsection 2.2.1 outlines the formal definition of the problem. Then, the following subsections present three different measurement techniques addressing this challenge.

2.2.1 Problem Formulation

The goal of the workcell calibration is to estimate the pose of an object from the workcell, such as the workpiece, relative to a known base coordinate system. For example, the relative pose $\boldsymbol{p}_0^w$ between the workpiece's frame at C_w and the base coordinate C_0 is to be estimated. The pose of the workpiece in the world coordinate system is given by the rigid transformation $\mathbf{T}_0^w$ as follows:

$$\boldsymbol{p}_0^w := \boldsymbol{T}_0^w. \tag{2.1}$$

The formulation of the problem is graphically represented in Fig. 2.1.

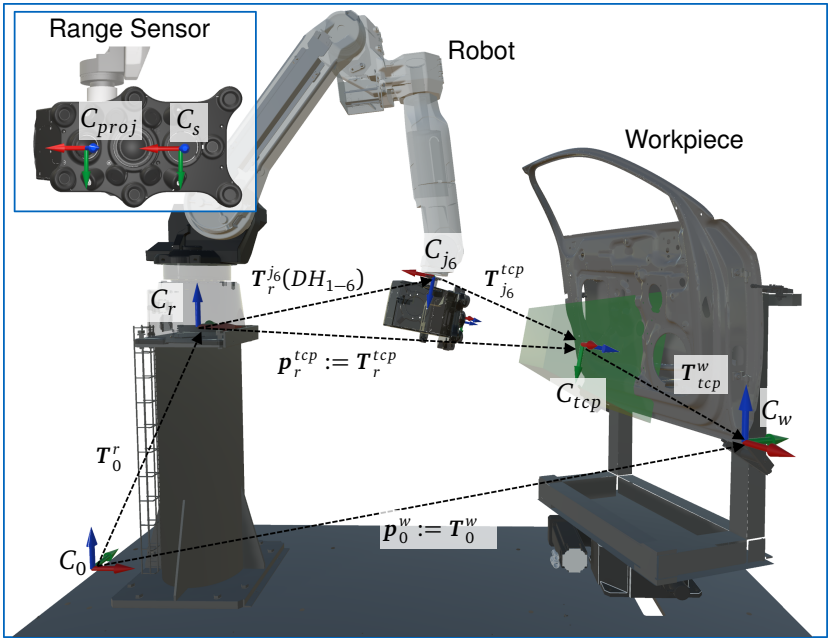


Figure 2.1: Overview of the kinematic model of a ROMS and the coordinate systems of its core components: base coordinate system C_0, robot C_r, sensor C_{tcp}, and workpieces C_w. The pose of the workpiece in the base coordinate system is denoted as follows: $\boldsymbol{p}_0^w := \boldsymbol{T}_0^w$.

A straightforward approach to estimate the rigid transformation given by Eq. 2.1 is to measure manually the distance between the two frames. However, in the context of ROMS, the rigid transformation can also be computed automatically

using the robot kinematic model. Assuming that the robot's kinematic model is known, the rigid transformation $\mathbf{T}_0^{tcp}$ between the base (C_0) and the TCP reference frame can be seamlessly estimated for each robot configuration (see Subsection A.1.1). For example, assuming that the transformation $\mathbf{T}_0^r$ between the base coordinate and the robot base is known, using the robot's kinematic model from Eq. A.1 and the transformation between sensor and workpiece $\mathbf{T}_{j_6}^{tcp}$, $\mathbf{T}_0^{tcp}$ is given:

$$\boldsymbol{T}_0^{tcp} = \boldsymbol{T}_0^r \cdot \boldsymbol{T}_r^{j_6}(DH_{1-6}) \cdot \boldsymbol{T}_{j_6}^{tcp}. \tag{2.2}$$

Given the rigid transformation $\mathbf{T}_w^{tcp}$ between the TCP and the workpiece, the pose of the workpiece in the world coordinate system is calculated as follows:

$$\boldsymbol{T}_0^w = \boldsymbol{T}_0^{tcp} \cdot \boldsymbol{T}_{tcp}^w. \tag{2.3}$$

Since the transformation $\mathbf{T}_0^{tcp}$ is known, the workpiece calibration can be seamlessly automated by calculating the transformation between the sensor's TCP and the workpiece $\mathbf{T}_w^{tcp}$. The following subsections introduce different approaches for estimating the kinematic relationship between the sensor's TCP and the workpiece frame.

2.2.2 Tactile Alignment

Tactile alignment is a common approach for aligning workpieces for machine tools and robot-based systems. The principle of this method consists of touching a workpiece's reference frame from different points with a robot's known reference frame, such as the edge of a tactile probe or the end-effector (BRECHER and WECK 2021, p. 594). Then, by approaching the same frame from different directions or using different reference frames and repeating the process, the position and orientation of the workpiece reference frame is estimated using a system of linear equations. BRECHER and WECK (2021, pp. 594–595) provide an overview of different strategies specifying different positions and characteristics used for tactile alignment. Although this approach is considered a standard alignment method in many applications due to its efficiency and simplicity, it generally requires an experienced operator to manually position the robot and avoid collisions.

2.2.3 2D Image based alignment

Identifying an object of interest within an unknown environment and estimating its position can be achieved through the utilization of 2D images captured by a camera. The fundamental research of this problem has its foundations in the face and pedestrian detection using image processing methods (SZELISKI 2022, p. 379; VIOLA and JONES 2001). However, academia and industry have extended the applicability of these methods to address object detection and pose

estimation tasks in industrial applications, e.g., bin-picking, quality assurance (DIRR et al. 2023; HARTL et al. 2021). This section presents a brief overview of different methods and variations addressing this task. A comprehensive overview of the latest research on 6D pose estimation can be found in DU et al. (2021), SHETTY et al. (2021), and ZHU et al. (2022).

Image-Based Object Detection

The initial step to estimate the position of an object is to robustly identify pixel patterns belonging to an object or its parts. Such patterns are referred to as *keypoint features* or *interest points* (SZELISKI 2022, p. 419). Once such reference points are selected, they are used to match objects with the same features in different images, ultimately helping to identify an object (BEYERER et al. 2016, p. 715).

Therefore, a fundamental problem in object detection is the adequate selection and synthesis of keypoint features. In the research field of classical image processing, different studies described mathematically different types of hand-designed features (e.g., points and patches, edges, curves and contours, straight lines) to identify image patterns (SZELISKI 2022, pp. 434–440). For example, object detection techniques based on hand-designed features (e.g., Haar-like features, scale invariant feature transform (SIFT), and histogram of oriented gradients (HOG)) have demonstrated solid robustness and accuracy within well-controlled environments. However, most of these techniques fail to capture mid- and high-level correlations (e.g., edges, intersections, or objects parts) composed by multiple features affecting its scaling and generalization performance (DENG 2014, p. 321).

On the contrary, in recent years deep learning models (see Subsection A.1.3.2) have exhibited a growing capability to tackle two difficult problems linked to the employment of traditional hand-crafted feature techniques. First, the design of more informative features for visual tasks can be automated (KRIZHEVSKY et al. 2017). Second, the architecture of *artificial neural networks* (ANNs) (see Subsection A.1.3.2) allows the synthesis of low- and high-level features, contributing to achieving better performance in terms of accuracy, generalization, biased data, and scalability (GIRSHICK et al. 2014; KRIZHEVSKY et al. 2017). For these reasons, the latest studies over the last decade have focused primarily on developing deep learning models for object detection (LIU et al. 2016; REDMON et al. 2016; REN et al. 2017) and segmentation[2] (BOLYA et al. 2019; HE et al. 2017; PINHEIRO et al. 2015).

[2]Segmentation refers to the process of dividing an image into regions and calculating a precise boundary to adjacent pixels so that each segment corresponds to an object or parts of it. (BISHOP 2006, p. 428; SZELISKI 2022, p. 387).

Pose Estimation

Once an object has been identified in an image, the next step considers the object's pose estimation. For instance, BRAUER (2014) and MÜLLER and ARENS (2010) used hand-designed features and a probabilistic–based approach to estimate the 3D pose of human bodies using person limbs as landmarks. Moreover, the results of BRAUER (2014, p. 212) demonstrated that the 3D pose of a human body could be estimated with a mean error up to 20 cm (distance of 2.9 m) using a camera projection model.

On the contrary, more recent studies have proposed *convolutional neural networks* (CNNs) (see Subsection A.1.3.2) to combine object detection and pose estimation. For example, KENDALL et al. (2015) presented one of the first models for 6D pose estimation using CNNs. In their study, the authors extend the CNN model proposed by SZEGEDY et al. (2015) to infer a detected object's position and orientation. Their model was evaluated for localization and mapping tasks for public buildings. The authors showed that their approach outperformed hand-designed feature detector approaches. Furthermore, KEHL et al. (2017) extended the object detection model of LIU et al. (2016) to estimate the 2D pose of an object. In the following step, the 6D pose is inferred based on a comparison of rendering hypotheses of 6D bounding boxes and a camera projection model to calculate the depth. Similarly, XIANG et al. (2018) proposed a camera projection model to estimate the depth of previously detected and segmented objects using a CNN. The authors introduce a CNN architecture designed for object detection, segmentation, 2D object center, and rotation regression. However, the regression of the rotation showed to be sensitive to ambiguities and convergence problems.

Synthetic Datasets for Model Training

Training supervised image-based *machine learning* (ML) models requires a large and labeled dataset to achieve reasonable performance. In particular, labeling images, i.e., estimation of bounding boxes, segmentation masks, and orientation labels, requires significant effort and expertise (HODAN et al. 2017). Therefore, some of the revised works using data-based approaches proposed to train their models using synthetic datasets (KEHL et al. 2017; SUNDERMEYER et al. 2018; XIANG et al. 2018). For instance, the work of TREMBLAY et al. (2018b) suggests a series of domain randomization techniques (background images, lighting conditions, different viewpoints) for generating synthetic datasets and reducing the training effort without sacrificing performance. The use of such methods has shown great potential and awakened great interest in recent years, so much so that even an industrial consortium of different project partners (BMW, NVIDIA, Microsoft, et al.) created SORDI[3], a public synthetic object dataset, optimized for training and validation of CNNs.

[3] Synthetic Object Recognition Dataset for Industry. https://sordi.ai/

2.2.4 Point cloud based alignment

2D imaged-based alignment techniques have shown that the 6D pose of an object can be robustly estimated from a single image to an accuracy of a few centimeters (BRACHMANN et al. 2016; HINTERSTOISSER et al. 2013; HODAŇ et al. 2016; ZHU et al. 2022). However, in many cases the alignment accuracy reached by these approaches is not sufficient. Therefore, many of the studies following a 2D image based alignment extended their methods to integrate extra depth information using point clouds to enhance the accuracy of their methods (KEHL et al. 2017; SUNDERMEYER et al. 2018; XIANG et al. 2018). For this purpose, local registration algorithms, such as the *iterative closest point* (ICP) algorithm of BESL and MCKAY (1992), take over the fine alignment and use the image-based rough pose estimation as an initial value. By integrating depth-information these studies demonstrated that the pose estimation accuracy could be considerably improved.

Although range imaging approaches allow for non-contact and accurate alignment, any detection method that relies on depth information still requires the sensor to be positioned close enough so the object lies within the sensor's measurement volume (see Section 1.2.1.2). For this reason, range image alignment, similar to tactile calibration, is also difficult to automate and requires a previous spatial analysis of the scene. In addition, the robust acquisition of depth information using optical sensors is more sensitive to lighting conditions and reflective material properties (BRACHMANN et al. 2016) (see Subsection 2.4).

2.2.5 Summary and Need for Action

Within this section three different techniques were introduced for estimating the pose of a reference object, using a tactile probe, a 2D image, or a 3D point cloud. First, tactile calibration (see Subsection 2.2.2) was identified as a highly accurate and common approach for programming machine tools and robot-based systems. Moreover, workcell calibration based on point clouds was also recognized as an accurate approach used by production systems that incorporate a 3D sensor (see Subsection 2.2.4).

Although tactile and point cloud calibration techniques are considered highly accurate and established approaches for calibrating workcells, both require significant human intervention for programming a routine to prevent collisions and ensure the quality of the calibration. For this reason, Subsection 2.2.3 presented an overview of approaches that allow object recognition using 2D images. Image-based calibration allows objects to be roughly detected from long and safe working distances, making it a viable approach for automating calibration tasks.

Furthermore, recent research focusing on image-based object detection indicates that data-driven models, such as CNNs, have established as the current state-of-the-art approach for solving these tasks (BRACHMANN et al. 2016; KEHL et al.

2017; KENDALL et al. 2015; SUNDERMEYER et al. 2018; XIANG et al. 2018). However, to achieve satisfactory performance, the models must first be trained on large datasets (GOODFELLOW et al. 2016). The generation of such datasets is still considered a significant resource-intensive task, making use of data-based approaches impractical for industrial applications.

Despite the well-founded solutions proposed to detect and estimate the pose of an object of interest based on images, end-to-end solutions have not yet to been established within the industry or academia. One reason for this discrepancy is that the efficacy of data-driven models depends heavily on the attributes of the training datasets (such as their size, quality, and variability) and the limitations of the particular applications (ZHOU and WU 2011). Furthermore, the current CNNs architectures require more lightweight designs to reduce the problem complexity and training effort (ZHU et al. 2022).

2.3 Viewpoint Planning Problem

Having an aligned kinematic model of the workcell, the next step in programming ROMSs deals with the generation of a viewpoint plan that satisfies the targeted inspection task. In academia, this generic challenge is commonly known as the *Viewpoint Planning Problem* (VPP). Subsection 2.3.1 introduces its formal definition.

The research in this field varies based on the defined system constraints and how domain knowledge (such as robotics, optics, metrology) is abstracted and formalized to address the VPP. Hence, Subsection 2.3.2 provides first an overview of commonly acknowledged viewpoint constraints that have been taken into account within the scope of metrological tasks and ROMS. Then, Subsections 2.3.3 and 2.3.4 provide an overview of studies that utilize prior knowledge and those that do not necessitate to solve the VPP.

The present subsection offers an overview of the related research in the scope of this thesis, an exhaustive overview of the overall progress, challenges, and further applications of the VPP is provided by the surveys of CHEN et al. (2011), KRITTER et al. (2019), MAVRINAC and CHEN (2013), PEUZIN-JUBERT et al. (2021), SCOTT et al. (2003), and TARABANIS et al. (1995a).

2.3.1 Problem Formulation

The VPP and related work addressing its resolution is better understood by considering the following minimal problem formulation based on the study of TARBOX and GOTTSCHLICH (1995):

"What is the minimum number k of viewpoints v required to acquire a given set of features F?"

Hence, a metrology task can be considered fulfilled when a viewpoint plan holding a finite number of k viewpoints guarantees the acquisition of all features F. In addition, it must be considered that a valid viewpoint must satisfy a defined set of viewpoint constraints $\widetilde{C}$.

Although the concept of viewpoints is used in the majority of the related research, there seems to be no established formal definition of a viewpoint v among them. Hence, this thesis considers a feature-centered formulation and defines a viewpoint as being a triple of following elements: a sensor pose $\boldsymbol{p}_s \in SE(3)$ to acquire a subset of features $G \subseteq F$ considering a set of viewpoint constraints $\widetilde{C}$:

$$v := (\boldsymbol{p}_s, G, \widetilde{C}). \tag{2.4}$$

2.3.2 Viewpoint Constraints

The VPP is still considered an open problem in the automation of various applications that depend on the computation of valid viewpoints, such as camera surveillance, scene exploration, object detection, visual servoing, object reconstruction, image-based inspection, robot calibration, and mobile navigation (CHEN et al. 2011; PEUZIN-JUBERT et al. 2021). Since viewpoint constraints vary within applications, some studies dealing with automated inspection (CHEN and LI 2004; SCOTT 2009; TEKOUO MOUTCHIHO 2012), explicitly defined different viewpoint constraints to first assess the validity of the generated viewpoints in the context of feature-based inspection tasks using ROMSs.

The following list presents the definitions of eight viewpoint constraints considered in this thesis. Some of these constraints have partial overlap with those addressed in existing literature. Figure 2.2 shows a simplified 2D representation of some of these constraints.

$\widetilde{C}_1$ Measurement Volume: The working space of range sensors is represented by their measurement volume. The measurement volume is described by different sensor parameters, e.g., the horizontal and vertical *field of view* (FOV) angles, resolution, focal length (COWAN and KOVESI 1988), see Subsection A.1.2.2. Therefore, the measurement volume combines multiple constraints and is regarded as a crucial and restrictive limiting factor in several applications. This constraint is fulfilled when the targeted surface of a workpiece lies inside the measurement volume (ABRAMS and ALLEN 1992).

$\widetilde{C}_2$ Incidence Angle: When selecting the sensor position for detecting a surface point, additional conditions regarding its orientation must be fulfilled. For most imaging sensors, a maximum incidence angle that describes the feature's normal and the sensor's optical axis must not exceed a threshold to ensure the acquisition of the surface. (SCOTT 2002, p. 105; REED 1998, p. 33)

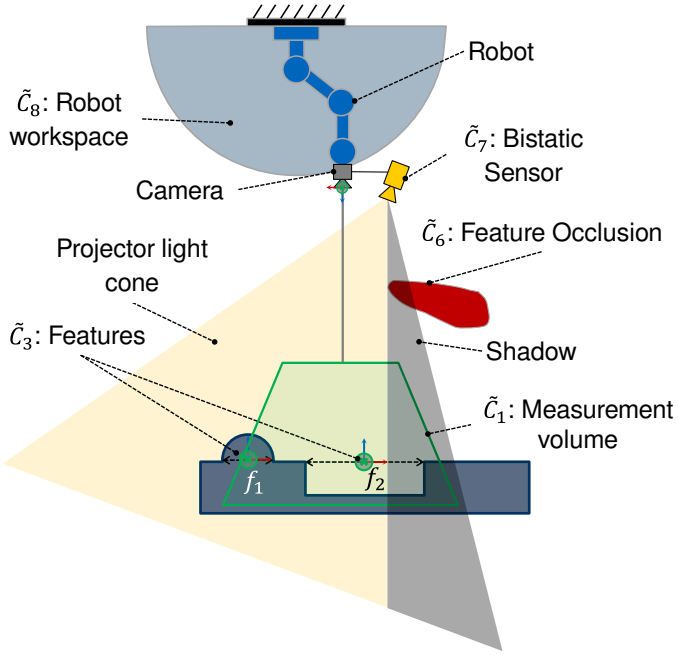


Figure 2.2: Simplified 2D graphical representation of some viewpoint constraints. The current viewpoint is invalid because some constraints ($\widetilde{C}$) are not satisfied. For example, the entire geometry of feature f_1 is not completely within the measurement volume ($\widetilde{C}_1$, $\widetilde{C}_3$). The occluding object prevents feature f_2 from being fully illuminated by the projector ($\widetilde{C}_6$, $\widetilde{C}_7$).

$\widetilde{C}_3$ Feature Geometry: In the context of inspection applications, it is important that all surface points of a feature are robustly acquired with a single measurement (CHEN and LI 2004). Therefore, this constraint is considered an extension of the first constraint and is satisfied if the entire feature geometry lies within the measurement volume and can be acquired with a single viewpoint.

$\widetilde{C}_4$ Kinematic Errors: In the context of real applications, system modeling inevitably involves discrepancies between virtual and real-world models (HIRZINGER et al. 1994; TEKOUO MOUTCHIHO 2012, p. 31). Therefore, any factor that affects the accuracy of the nominal position of the sensor and impact the validity of a viewpoint must be considered. Such factors include kinematic modeling errors of the robot, sensor, and workpiece, or the positional accuracy of all manipulators (SCOTT 2002, pp. 170–174).

$\widetilde{C}_5$ Sensor Accuracy: Given that the sensor's accuracy and sampling density may vary within its measurement volume, it is necessary to consider this characteristic constraint for the viewpoint planning. For instance, SCOTT (2009) presented a model for quantifying these variables within

the sensor's measurement volume. In addition, PARK et al. (2006) showed that there may be quality differences of the sensor at different acquisition planes and proposed a look-up table to represent them. Therefore, a constraint representing the sensor accuracy must be considered when the sensor accuracy can be quantified.

$\widetilde{C}_6$ Feature Occlusion: A viewpoint is considered valid if a free line of sight exists between the sensor and the feature including its whole geometry (see third constraint). For this reason, a free occlusion constraint must be regarded to guarantee that no object blocks the sensor's visibility. (CHEN and LI 2004; REED 1998)

$\widetilde{C}_7$ Bistatic Sensor: Recalling the bistatic nature of range sensors, two imaging devices (two cameras or one camera and one projector) are required to compute a point cloud (SCOTT 2002; TEKOUO MOUTCHIHO 2012). For this reason, all viewpoint constraints must be fulfilled for all cameras or projectors of a range sensor at the same time.

$\widetilde{C}_8$ Robot Workspace: Since the sensor is attached to the robot, additional constraints must be fulfilled to ensure a viewpoint's validity. A viewpoint is valid if this is reachable by the robot, i.e., it lies within the robot workspace (CHEN and LI 2004).

2.3.3 Model-Based Planning

In the context of metrological applications, it is assumed that the workpiece position and the nominal position and orientation of the features are known. This enables exploitation of this information for planning more effective viewpoints. All techniques, which utilize *a priori* knowledge about the workpiece for viewpoint planning fall under the category of model-based approaches (PEUZIN-JUBERT et al. 2021; SCOTT 2002).

Furthermore, the works following a model-based approach are differentiated on how the features are modeled and how the components and constraints of a ROMS are abstracted and formalized to solve the VPP. Most model-based viewpoint planning methods can broadly divided into *synthesis* and *sampling-based* approaches (TARABANIS et al. 1995b).

Synthesis

Synthesis approaches aim to model viewpoint constraints in the form of continuous or discrete solution spaces before searching for an optimal viewpoint. Techniques aiming at characterizing search spaces[4] have been proposed since the first studies dealing with the VPP. Such a formulation has the advantage

[4](related terms: viewpoint space, visibility map, visibility matrix, visibility volumes, imaging space, scannability frustum, configuration space, visual hull, search space)

of providing a straightforward understanding and spatial interpretation of the problem.

The first studies that considered the characterization of a continuous solution space in the Euclidean space $\mathbb{R}^3$ can be attributed to the publication of COWAN and KOVESI (1988). In their work, they introduced a model-based method for 2D sensors, which synthesized analytical relationships to characterize a handful of constraints geometrically: resolution, focus, field of view, visibility, view angle. Furthermore, COWAN and BERGMAN (1989) introduced constraints for placement of a lighting source. Based on the analytical findings provided by the previous work of COWAN and BERGMAN (1989) and TARABANIS et al. (1995a) introduced a model-based sensor planning system, which considered occlusion-free regions. In addition, the authors posed the problem in an optimization setting using objective functions to find valid viewpoints within the occlusion-free space that meet imaging constraints. The planning system proposed by TARABANIS et al. (1995a) was extended by ABRAMS et al. (1999) for its use with industrial robots and moving objects. Their study addressed the drawbacks (nonlinearity and convergence guarantee) of the optimization algorithms and opted to characterize 3D search spaces for the sensor's resolution, the field of view, and the robot's workspace. Although the authors could not synthesize every constraint in the Euclidean space, they confirmed the benefits of solving the problem in $\mathbb{R}^3$ instead of optimizing equations for finding suitable viewpoints. Furthermore, in a series of publications REED and ALLEN (2000) and REED (1998) extended some of the models introduced by TARABANIS et al. (1995a) and addressed the characterization of a search space in $\mathbb{R}^3$ for range sensors, which integrates imaging, occlusion, and workspace constraints.

Another line of research was followed by the works of TARBOX and GOTTSCHLICH (1994, 1995) which proposed the synthesis of a discretized search space based on *visibility matrices*. These matrices are used to map the visibility between the solution space spanned by the imaging parameters of the sensor and the discretized surface of an object. In their works, TARBOX and GOTTSCHLICH presented different algorithmic variants based on the concept of visibility matrices to solve the VPP in the context of inspection tasks. Similarly, SCOTT (2002, 2009) extended the concept of visibility matrices by considering additional constraints, such as occlusion, sensor accuracy, and variation of incident angles. Moreover, recent studies demonstrate the benefits of visibility matrices in encoding information between the surface of an object and its related viewpoints (GOSPODNETIĆ et al. 2022; GRONLE and OSTEN 2016; JING et al. 2017; MOSBACH et al. 2021).

In the context of space discretization and feature-driven approaches, the publications of PITO (1999) and TRUCCO et al. (1997) suggest characterizing the solution space as a sphere with all viewpoints on its surface. This approach allows to simplify the 6D sensor positioning problem to a 2D optimization problem with a fixed orientation. Similarly, STÖSSEL et al. (2004) and ELLENRIEDER et al. (2005a) introduced the concept of visibility maps to encode feature visibility of a sphere in a matrix. RAFFAELI et al. (2013) and KOUTECKÝ et al. (2016) considered variations of this approach for their viewpoint planning systems.

Furthermore, other relevant works that focused on laser scanners, (DERIGENT et al. 2006; LEE and PARK 2000; TEKOUO MOUTCHIHO 2012), also considered the synthesization of a search space before searching for feasible solutions.

Sampling-Based

Sampling techniques introduce objective functions to model constraints. The satisfiability of each viewpoint constraint is then evaluated individually by sampling the search space using metaheuristic optimization algorithms, such as simulated annealing or evolutionary algorithms (CHEN and LI 2004; ERDEM and SCLAROFF 2006; GLORIEUX et al. 2020; GONZÁLEZ-BANOS 2001; MAVRINAC et al. 2015). To a certain extent, many studies adopting a synthesis approach also utilize sampling techniques to evaluate residual constraints on viewpoints that cannot be synthesized analytically.

2.3.4 Non-Model Based Planning

In contrast, *non-model based* approaches require no *a priori* knowledge; the object's location and form is unknown to the planning system. In this case, online exploratory techniques based on the captured data are used to compute the next viewpoint (KRIEGEL et al. 2015; LAURI et al. 2020). Most of the strategies proposed focus on reconstruction tasks and address the problem as the *next-best-view* planning problem. Recalling the outlined coordinate metrological task (see Subsection 1.2.2), it is assumed that the features and their location are known *a priori*. Hence, related techniques neglecting this information fall outside the scope of this thesis and are not further discussed.

2.3.5 Summary and Need for Action

An overview of different approaches addressing the VPP, i.e., the computation of the minimum number of valid viewpoints to complete a vision task were presented in this section. Although numerous sophisticated solutions have been proposed to tackle the VPP, their transferability, efficiency, and effectiveness are often restricted by the limitations of the individual vision systems and the specific tasks targeted. Based on the literature review, the present thesis did not find an appropriate approach to solve the VPP for metrology tasks and satisfy all considered viewpoint constraints considered in Subsection 2.3.2. Specifically, two pressing issues related to this research gap were identified.

Problem Formulation and Consistent Characterization of Viewpoint Constraints

The first point deals with the appropriate formulation of the VPP itself. On the one hand, there exists exists a literature trend, which poses the VPP as an

optimization problem (Beasley and Chu 1996; Chen and Li 2004; Erdem and Sclaroff 2006; Glorieux et al. 2020; Kaba et al. 2017; Mavrinac and Chen 2013; Mittal and Davis 2007; Scott 2009). These studies model each viewpoint constraint as objective functions. Then, optimal and valid viewpoints are found using meta-heuristic optimization algorithms, e.g., simulated annealing or evolutionary algorithms.

On the other hand, another line of work takes a more analytical approach and initially aims to spatially define viewpoint limitations as solution spaces (Kouteckỳ et al. 2016; Mosbach et al. 2021; Scott 2002; Tarbox and Gottschlich 1994; Tekouo Moutchiho 2012). In the second step, an optimization algorithm selects optimal viewpoints within the synthesized solution spaces. Viewpoint constraints that could not be analytically formalized are individually assessed using binary functions.

In general, most of the revised works follow different approaches to synthesize viewpoint constraints. Thus, the adaptability of the solutions and their effectiveness depend strongly on the fundamental methods used to model these. This inconsistency has led to the development of rigid, tailor-made solutions, making their generalization for different use-cases or viewpoint constraints difficult. Therefore, this thesis considers that a modular formulation of the VPP and a consistent modeling of the viewpoint constraints are key elements towards more efficient and generic solutions.

Compensation of Modeling Uncertainties

In the context of industrial applications, a second fundamental pressing issue, which few researchers addressed, is the consideration of model uncertainties such as the positioning error of the system or sensor inaccuracies (Scott 2002, p. 212; Tekouo Moutchiho 2012, p. 133). System modeling inevitably involves discrepancies between virtual and real-world models, especially in dynamically changing environments. For this reason, some of the first studies dealing with the VPP, followed an analytical approach and suggested modeling solution spaces to formalize the spatial influence of different viewpoint constraints (Cowan and Kovesi 1988; Reed 1998; Tarabanis and Tsai 1991; Tarbox and Gottschlich 1994). Solution spaces can be used to compensate for modeling uncertainties in real applications (Scott 2002; Tarbox and Gottschlich 1995). For example, if a selected viewpoint results invalid, an alternative viewpoint is seamlessly selected from the solution space.

More recent works have continued to assess a viewpoint's validity using similar approaches. Some prominent examples are the works of (Gospodnetić et al. 2022; Gronle and Osten 2016; Jing et al. 2017; Kouteckỳ et al. 2016; Mosbach et al. 2021), who followed the concept of Scott (2002) and Tarbox and Gottschlich (1995) to discretize the viewpoint space and the object surface space based on visibility matrices. Although these works have steadily demonstrated the effectiveness of visibility matrices to solve the VPP, the main weakness of these approaches is related to the inherently limited storage capacity

and the computational efficiency and inaccuracy associated with discretizing the object's surface and the viewpoint space. In addition, for large-scale metrology applications that consider large workspaces, discretizing solution spaces could lead to computationally inefficient solutions. Furthermore, most of these works reduce the problem dimensionality for sensor positioning using a fixed working distance, limiting the intrinsic flexibility of ROMS and the overall solution space.

Although the characterization of solution spaces is a sound approach for computing valid viewpoints and addressing modeling uncertainties, the current research lacks a consistent modeling framework for characterizing all viewpoint constraints considered in the scope of this thesis. Therefore, the spatial and explicit characterization of all considered viewpoint constraints is a fundamental unresolved issue required to tackle the VPP.

2.4 Measurement Parameterization

Measuring the 3D shape of reflective surfaces is a challenging task for optical metrology researchers due to drastic variations in intensity responses. In the last two decades, numerous studies have proposed different techniques to parameterize sensors and obtain suitable image exposure. The objective is to achieve a uniform image intensity, preventing over- or under-exposure of pixels.

To understand the complexity of acquiring an adequate image exposure, first, the problem's formal definition is outlined in Subsection 2.4.1. Then, the following subsections present various techniques proposed in the last two decades to optimize image exposure for active optical range sensors.

2.4.1 Problem Formulation

The image exposure, which is defined as the amount of light that reaches the camera sensors, determines the intensity of the resulting image and depends on several influencing factors, e.g., sensor focal length, lens aperture (exposure time), the light intensity of the projector, the surface reflectivity, external lighting sources (NAYAR et al. 1991; ELLENRIEDER et al. 2005a; EKSTRAND and ZHANG 2011; ZHANG 2020; GÜHRING 2002; SZELISKI 2022, p. 611). For instance, in monochrome cameras, the image intensity is commonly quantified as the gray level of an eight-bit scale $[0-255] \in I_s$, with 0 corresponding to a black pixel and 255 corresponding to white (GÜHRING 2002, p. 37).

EKSTRAND and ZHANG (2011) proposed a simplified linear model to calculate the average image intensity $I_s^{i,j}$ for each pixel x_i, y_j given the sensor's exposure time t_{exp} and a camera sensitivity γ_s:

$$I_s^{i,j} \sim t_{exp} \cdot \gamma_s \cdot [L_a^{i,j} + \rho_o \cdot (L_o^{i,j} + L_p^{i,j})] + \sigma_s. \tag{2.5}$$

The local average intensity for a surface area corresponding to a set of pixels depends on the following factors (EKSTRAND and ZHANG 2011):

- the direct ambient light intensity (L_a) going directly into the sensor,
- the ambient light intensity (L_o) reflected by the object with a surface reflectivity of ρ_s,
- the projected light intensity of the active source (L_p) reflected by the object with a surface reflectivity of ρ_s, and
- the sensor noise σ_s.

Furthermore, the studies of EKSTRAND and ZHANG (2011), FENG et al. (2014), and ZHANG (2020) assume that in the context of active sensing, the projector's light is generally more dominant than the ambient light and the resulting light reflected by the object. Therefore, in such cases the ambient light can be neglected yielding the simplification of Eq. 2.5:

$$I_s^{i,j} \sim t_{exp} \cdot \gamma_s \cdot \rho \cdot L_p + \sigma_s. \tag{2.6}$$

The model of EKSTRAND and ZHANG (2011) showed to be a valid approximation for static measurement configurations, which considered a constant working distance to the surface and a constant incidence angle between the lighting direction and the surface normal. However, these conditions do not fully hold for applications with specular objects requiring a dynamic sensor positioning.

In contrast to ideal diffuse objects, which reflect light uniformly in all directions (Lambertian reflection), specular objects scatter light depending on the incoming light direction (SZELISKI 2022, p. 80; GÜHRING 2002, pp. 80–82), see Fig. 2.3. For this reason, Eq. 2.6 must be extended for dynamic configurations by taking into the account the distance to the image plane z_s and the incidence angle ϕ_s (GÜHRING 2002, p. 81; COWAN and MODAYUR 1993):

$$I_s^{i,j} \sim t_{exp} \cdot \gamma_s \cdot \rho \cdot \frac{1}{z_s^2} \cdot \cos(\phi_s) \cdot L_p + \sigma_s. \tag{2.7}$$

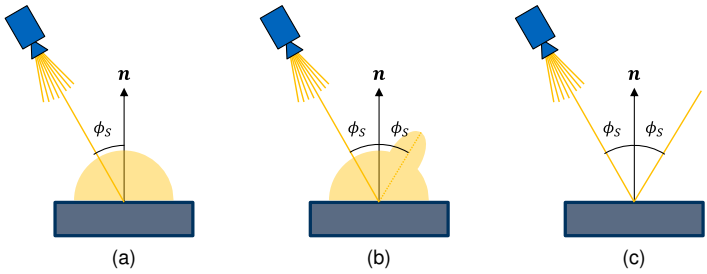


Figure 2.3: Reflection models depending on the incidence angle ϕ_s and surface reflectivity: a) ideal diffuse, b) directional diffuse, and c) ideal specular according to (GÜHRING 2002, p. 80)

The parameters of Eq. 2.7, such as the exposure time, projected light intensity, angle of incidence, and sensor distance, can be adjusted to control the image intensity range and obtain an appropriate exposure. In the subsequent subsections, various approaches are presented that aim to achieve optimal image intensity by modifying or controlling these parameters.

2.4.2 Approaches for Measurement Optimization

Spray Coating

A common method to achieve uniform image exposure is to mat the surface of the target by applying an anti-reflective material such as chalk or titanium (PALOUSEK et al. 2015). The aim of this procedure is to modify the acquisition surface properties making it more diffuse. This ensures that the reflected light by the object is distributed more homogeneously and is independent of the light source direction, as shown in Fig. 2.3a. Although this technique is still considered very effective in preventing overexposure of the camera image, this approach also has drawbacks, e.g., not all workpieces can be coated, the process usually requires manual pre- and post-processing steps (coating and cleaning), and the measurement accuracy is affected by the powder thickness and its distribution homogeneity (PALOUSEK et al. 2015; SHI et al. 2006).

High Dynamic Range

High dynamic range (HDR) techniques are considered a state-of-the-art approach for optimizing image exposure established within many commercial and industrial optical sensor systems (SZELISKI 2022, p. 609; BEYERER et al. 2016, p. 213). The core idea consist of merging multiple images acquired with different exposures to optimize the intensity of individual pixels or patches. Producing a series of images with multiple light intensities can be achieved mainly in two ways: camera exposure control and projector fringe pattern modulation (LIN et al. 2017b).

For example, ZHANG and YAU (2009) study utilizes an HDR method to create a pixel-by-pixel image by selecting the best illuminated pixels from a set of fringe images captured with varied exposure times. Moreover, FENG et al. (2014) suggested an empirical approach to analyze first the exposure of individual image regions using an initial exposure. Based on these findings, in the second step, multiple exposure times are then calculated and the corresponding images are merged to obtain an ideal exposure. Furthermore, SONG et al. (2017) investigated the acquisition of specular objects and proposed a pixel-wise fusion algorithm to combine different images taken with different exposure times. More recently, ZHENG et al. (2019) used a color camera for producing different intensities with a single fringe projection. The study proposes a HDR method, which fuses the images of the three camera channels (red, green, and blue) to generate

a well-exposed image. Although all of these works showed promising results for obtaining high-density point clouds in applications with a high dynamic range of surface reflectivity, the number and selection of exposure times were chosen empirically and their estimation was not discussed further.

Besides camera exposure control, there is an alternative research line that utilizes fringe pattern modulation to obtain different images with a high range of intensity levels. For example, JIANG et al. (2012) and SKOCAJ and LEONARDIS (2000) select empirically projector intensities to acquire multiple images with different intensities and fused them to obtain an optimized HDR image. WADDINGTON and KOFMAN (2010, 2014b) proposed a technique to adjust the projecting fringe pattern based on the maximum input gray level (MIGL) of the fringe images before reaching the image saturation. This value is estimated empirically beforehand and depends on the working distance of the sensor. In a later publication, WADDINGTON and KOFMAN (2014a) showed how this approach is used to generate composite HDR images with an optimized global exposure. Furthermore, LI and KOFMAN (2014) followed a similar approach, however, their study concentrated on estimating the MIGL locally for individual regions of an image. Moreover, ZHAO et al. (2014) addressed the disadvantages of methods using projectors in terms of time efficiency. By integrating more LED chips and using a modulation control of the light, the authors demonstrated that the acquisition speed can be significantly increased.

In the context of active range sensing, HDR techniques have been demonstrated to be suitable for surface measurement of non-diffuse objects. More detailed insight into the various HDR approaches and their applications are given in the review papers of FENG et al. (2018) and LIN et al. (2017a).

Multiple Cameras

Taking into account the advantages of HDR methods, an alternative way considered by other studies to obtain a high range of image intensities is to consider redundant measurement configurations comprising multiple cameras. In contrast to HDR techniques, multiple images can be acquired simultaneously, reducing the time required to acquire each image. LIU et al. (2011) proposed a dual-camera vision system, where the depth information of a surface point can be calculated from any of both cameras. Their study shows that by combining a multiple exposure method with a technique to merge the fringe images from both cameras, a better quality of the point cloud is obtained. A dual-camera setup was also proposed by FENG et al. (2017). The authors showed that saturated pixels from one camera can be compensated by the second camera at a different angle. The proposed system showed to be effective for compensating over- and underexposure areas with shiny surfaces.

Analytical Modeling

Another way of optimizing the exposure of an image is to address this problem from an analytical perspective. For instance, ELLENRIEDER et al. (2005b) and KOUTECKÝ et al. (2016) proposed an offline viewpoint planning system that estimates an adequate sensor's exposure time to acquire the surface of specular objects based on the sensor position and the reflectance simulation model framework of NAYAR et al. (1991). The study of KOUTECKÝ et al. (2016) presents a comprehensive validation using industrial metal sheets, where the resulting point cloud of the acquired surface area could be estimated with a deviation range between 13 % to 26 %.

In addition, the studies of EKSTRAND and ZHANG (2011) and ZHANG (2020) introduced an autoexposure technique considering a linear approximation of the image intensity based on Eq. 2.6. EKSTRAND and ZHANG (2011) proposed to pre-analyse a series of fringe images acquired with different exposure times to determine a globally optimized exposure time. ZHANG (2020) enhanced this technique by using only a single image and extended the framework for computing multiple exposure times that can be applied to HDR-based methods. Both studies show that the approximated linear model is valid and provides satisfactory results for static environments with a fixed sensor position and orientation.

Sensor Modifications

Another area of research focuses on improving image exposure through additional modifications of the sensor. GUPTA et al. (2011) studied the influence of fringe patterns on the image intensity of individual pixels. Their work suggested an alternative set of modified fringe patterns to compensate for some global illumination effects, e.g., inter-reflections or diffusion. BABAIE et al. (2015) adjust recursively pixel intensities of projected fringe patterns by a feedback control loop, which analyses the current image intensities. Similarly, LIN et al. (2017b) investigated an optimized projecting pattern based on a preliminary analysis of multiple fringe images generated with different gray-level intensities. Moreover, FENG et al. (2014) and SALAHIEH et al. (2014) showed that linear polarizers are effective in avoiding over-saturation of images.

Surface Reconstruction

Another approach to deal with reflections is to first identify the areas where specular reflections occur, and in the next step apply a region-filling method to reconstruct overly saturated pixels as proposed by SUN et al. (2017). However, reconstruction techniques are not suitable for accurately evaluating the acquired surface in quality control tasks, as they tend to distort the measurement.

2.4.3 Summary and Need for Action

Obtaining an adequate camera exposure is critical to ensuring robust and reliable sampling of an object's surface. Managing the camera exposure effectively, particularly for high-reflective surfaces, is a well-known challenge for optical range imaging sensors. This section provided an overview of different approaches to addressing this problem.

One of the most common and pragmatic methods to counteract inhomogeneous camera exposure with high-reflective materials is to matte-coat the surface object (cf. Subsection 2.4.2). However, this process is difficult to automate and the measurement accuracy may be affected by the coating (PALOUSEK et al. 2015; SHI et al. 2006). Furthermore, Subsection 2.4.2 provided an overview of HDR techniques that combine multiple images to produce a well-exposed image. Although HDR techniques have proven to be robust solutions, the usability of such methods is impractical for end users using commercial or industrial active systems if the vendor does not integrate this functionality. Commercial sensors typically do not allow for the necessary hardware or software modifications to automate this process. The majority of approaches focus on solving the problem using multiple empirically selected exposures. However, the fundamental problem of selecting an optimal exposure time falls on a second plane remaining unaddressed ZHANG (2020).

For this reason, ZHANG (2020) proposed a global exposure time optimization using a single captured stripe image and an approximated linear model of the exposure time and the image intensity. Although EKSTRAND and ZHANG (2011) and ZHANG (2020) extensively validated the validity and effectiveness of the linear approximation, they assumed a static measurement configuration. This configuration is limiting and insufficient for dynamic sensor positioning applications that use an ROMS and consider different acquisition depths and incidence angles. Furthermore, in real applications, a nonlinear response between the light source and the pixel intensity value at the sensor is to be expected (SZELISKI 2022, p. 611).

Furthermore, ELLENRIEDER et al. (2005a) and KOUTECKÝ et al. (2016) used analytical methods for estimating valid exposure times using the reflectance simulation model framework of NAYAR et al. (1991) and a 3D surface model of the probing object. Although KOUTECKÝ et al. (2016) demonstrated that the object's surface could be predicted with a deviation up to 26 %, the achieved performance suggests that the process could not be fully automated.

Moreover, in recent years, academia and industry have approached this problem from a different perspective by designing range sensors with more than one camera (FENG et al. 2017; LIU et al. 2011). However, similar to HDR techniques, this approach only partially solves the problem, as sensor positioning and exposure time selection remain open issues. In addition, merging images from different cameras increases the system's complexity.

Although several well-founded solutions for optimizing the camera exposure of

active 3D sensors have been proposed in academia and industry, the literature review undertaken for this thesis did not identify an effective approach that focuses on targeted and effective optimization of features within an image without modifying the hardware or software of the vision system. A local optimization has the potential to improve the effective acquisition of features and aim at the full automation of the measurement parameterization process. Therefore, achieving the highest measurement quality requires a comprehensive approach that considers the spatial relationships between the camera and the features.

Chapter 3

Solution Modules

The present chapter introduces the conceptualization of the solution modules to enhance the programming of *robot-based optical 3D measuring systems* (ROMSs) for performing metrology tasks. First, Section 3.1 introduces the overall concept of the present thesis and conceptualization of the solution modules. Then, Sections 3.1.1–3.1.3 present a summary of the implementation of the solution modules which is given by the PUBs1–4.

3.1 Overall Concept

The present research proposes a sensor planning system to increase the *level of automation* (LoA) of metrology tasks using ROMS. Fig. 3.1 depicts the main components of such a sensor planning system according to TARABANIS et al. (1995a) and the placement of this thesis's publications.

The proposed sensor planning system can be categorized as a hybrid-programming system. Although offline programming approaches have shown significant advantages for simplifying the creation of robot programs (cf. Subsection 2.1.2), there are rarely applications where a robot program created offline does not require a re-parameterization on site under real conditions (VOGL 2009, p. 20; BRECHER and WECK 2021, p. 596). This applies to metrology tasks that involve optical sensors. For this reason, this research focuses on metrology tasks utilizing optical sensors and proposes a hybrid approach that combines offline and online programming methods. Solution modules one and three are categorized as online methods as they incorporate techniques requiring the physical system. On the other hand, the second solution module utilizes a model of the ROMS to calculate viewpoints without requiring the physical system.

The subsequent subsections present an individual analysis of each solution module, which takes into account the conceptualization and necessary deliverables to address the stated *research questions* (RQs).

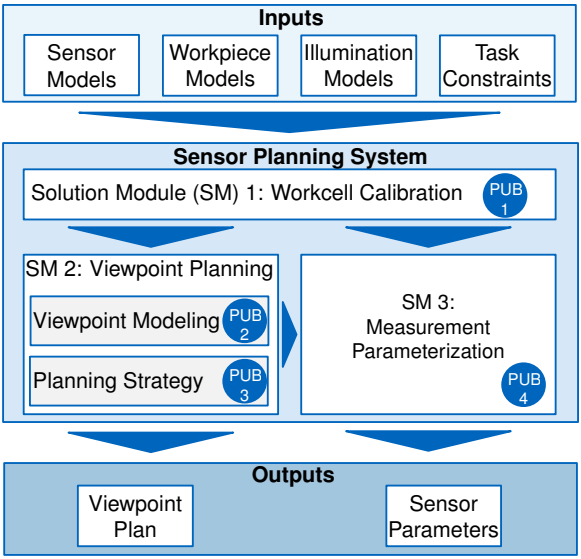


Figure 3.1: Overview of the solution modules of the present research being part of a sensor planning system as proposed by TARABANIS et al. (1995a)

3.1.1 Workcell Calibration

The first solution module deals with the calibration of the workcell kinematic model, i.e., the estimation of the pose of any workcell object (e.g., workpieces, fixtures) in a common coordinate system. In practice, this process is usually done manually. This subtask has the potential to be fully automated based on an optical alignment using the images of the ROMS's sensor. To conceptualize a solution to address this shortcoming, *RQ1* was proposed.

RQ1: How can the workcell calibration be automated by using images captured by the optical sensor of ROMSs?

3.1.1.1 Conceptualization of an Image-Based Workpiece Alignment

To address the posed RQ1, this thesis proposes an optical calibration of the workcell using a single camera image captured by the optical sensor of a ROMS from a safe distance. Based on the findings of the reviewed works, the present research argues that the sensor images, which provide a 2D representation of the scene, can be used to detect the workpiece inside the workcell and estimate its pose, thus solving the problem stated in Subsection 2.2.1. By estimating the workpiece's relative pose to the sensor (see Eq. 2.3), the workpiece's absolute pose can finally be calculated using the robot's kinematic model.

The design of a calibration strategy that combines different image processing methods promises to increase the LoA of the workcell calibration process. The following sub-modules have been identified as challenging tasks that need to be the focus of research to outline such a strategy.

Modular pose estimation strategy This thesis poses the estimation of an object from a single camera image as a multi-stage problem consisting of the following subproblems: *object detection and segmentation, position estimation,* and *orientation estimation*. The present research proposes a holistic strategy that combines different data-based and classical image processing methods to address the individual subproblems of the pose estimate problem. For example, *convolutional neural networks* (CNNs) (see Subsection A.1.3.2) have the potential to solve complex vision tasks such as object detection based on images within unknown environments. However, these models usually require large datasets to optimize multiple hyperparameters and achieve satisfactory performance. Therefore, a strategy which combines different image processing methods is to be outlined as a first step. For example, a straightforward camera projection model[1] can be applied to determine the location of an identified object in an image (BRAUER 2014, pp. 193–195), as opposed to a more sophisticated data-based approach.

Industrial compliant dataset generation Data-based models require a large dataset to train the models and ensure their performance. This requirement often conflicts with and hinders the use of data-based approaches within industrial applications. Therefore, a user-friendly and resource-effective solution to address this challenge is to be outlined within this solution module.

In the context of the digitization of production, it has become a common practice to have 3D *computer-aided design* (CAD) models of at least the key components and resources (e.g., workpieces, assemblies, fixtures) required for end-product manufacturing. This thesis proposes to exploit these 3D models to create synthetic datasets. Synthetic datasets can be generated using rendering[2] tools to synthesize photorealistic images that can be used for data-based model training (TREMBLAY et al. 2018a; ZHANG et al. 2017). The purpose of conceptualizing such a strategy is to assist end-users in minimizing the effort required to produce high-quality datasets, while simultaneously automating the entire process.

[1] A camera projection model refers to a simplified linear model commonly used to model a real imaging device. A standard model used in academia and industry is the *pinhole camera* model. The pinhole camera is defined by a set of intrinsic parameters that model the imaging characteristics (e.g., camera focal lengths, camera center coordinates, scaling factor) and the extrinsic parameters that represent the camera pose within the world coordinate system. (BEYERER et al. 2016, p. 103)

[2] Rendering refers to the process of synthesizing a digital image from a 3D scene (MOHAMMADIKAJI 2019, p. 65).

3.1.1.2 Solution Module Deliverable

The expected deliverable for this solution module is summarized as follows:

> Methodology for estimating the pose of an object of interest using a single image.

3.1.2 Viewpoint Planning

Based on an inspection plan and the calibrated kinematic model of the workcell, the second solution module addresses the computation of all the necessary viewpoints to acquire all features considering a set of constraints. The following RQ formalizes the problem and guides the conceptualization of this solution module:

> *RQ2: How do viewpoint constraints affect sensor positioning, and how can the minimum number of viewpoints required to capture all features be estimated?*

3.1.2.1 Conceptualization of a Viewpoint Planning Strategy based on Solution Spaces

To address the identified challenges associated with the *Viewpoint Planning Problem* (VPP) and both subquestions of RQ2, the present thesis proposes first to modularize the VPP into two subproblems, the *Viewpoint Generation Problem* (VGP) and the *Set Covering Problem* (SCP). The VGP deals only with the computation of valid viewpoints to acquire a single feature considering different viewpoint constraints. On the contrary, the SCP focuses on finding the minimum number of viewpoints necessary to acquire all features. This thesis considers that a well-formulated multi-stage solution has the greatest potential for reducing the overall complexity of the VPP and provides the opportunity to outline more generic and effective solutions for each subproblem. The subproblems and the approaches chosen to address both subproblems are explained in more detail below.

Explicit Characterization of Viewpoint Constraints as *feature-based constrained spaces* ($\mathcal{C}$-spaces) To formulate a generic and consistent approach for answering the first subquestion of RQ2 and addressing the VGP, this thesis proposes that any viewpoint constraint that affects the positioning of the sensor to capture a single feature is spatially modeled as a continuous solution space ideally in the special Euclidean $SE(3)$. Therefore, the present research introduces the concept of $\mathcal{C}$-spaces to denote such spaces. An abstract representation of individual $\mathcal{C}$-spaces for different features is depicted on the left side of Fig. 3.2. In addition, the same graphic illustrates the intrinsic advantage of $\mathcal{C}$-spaces for

dealing with modeling uncertainties. If the validity of a selected viewpoint is compromised, an alternative viewpoint can be seamlessly selected within the $\mathcal{C}$-space. In addition, if all viewpoint constraints can be modeled as $\mathcal{C}$-spaces, they can be combined with each other to span another solution space that guarantees the satisfaction of multiple constraints.

Viewpoint Planning Strategy based on *feature-cluster constrained spaces* ($^{G}\mathcal{C}$-spaces) The present research assumes that if there exists a $\mathcal{C}$-space for each feature, multiple spaces can be merged into a common $^{G}\mathcal{C}$-spaces for acquiring a group of features. The design of a holistic strategy that finds a potential group of features that can be acquired together with a common $^{G}\mathcal{C}$-space is the focus of the second subquestion of RQ2 and the basis for solving the SCP. A simplified representation of the problem and the proposed solution using a $^{G}\mathcal{C}$-space is shown on the right side of Fig. 3.2.

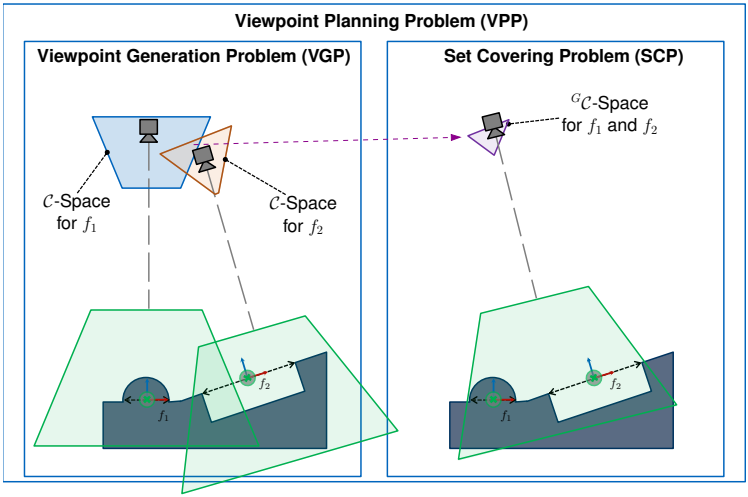


Figure 3.2: Modularization of the VPP and simplified representation of its subproblems, the VGP and the SCP. The VGP is solved based on $\mathcal{C}$-spaces, which characterize the solution space to acquire one feature while fulfilling different viewpoint constraints. The SCP is solved using $^{G}\mathcal{C}$-spaces, which span a solution space for capturing multiple features and simultaneously fulfill all viewpoint constraints.

3.1.2.2 Solution Module Deliverable

Having decomposed the problem, the expected deliverable for this solution module is summarized below.

Viewpoint planning framework for computing valid viewpoints that ensure the acquisition of multiple inspection features and can compensate for model uncertainties. All valid viewpoints satisfy all viewpoint constraints (see Subsection 2.3.2).

Followed by the proposed modularization of the problem, two publications were published to address the individual subproblems of the VPP and answer the subquestions of RQ2.

3.1.3 Measurement Parameterization

Automating the sensor parameterization process to achieve a dense, high-quality point cloud is essential for minimizing the manual labor involved in this task. Therefore, the following RQ3 was posed to guide the conceptualization of a solution to this challenge.

RQ3: How do image exposure and the spatial correlations between features and the sensor affect the successful acquisition of a point cloud, and how must these be considered for setting a proper exposure time?

3.1.3.1 Conceptualization of Data-Based Models for Prediction the Camera's Exposure Time and the Point Cloud Quality

Selecting an appropriate exposure time remains a time-consuming manual task, resulting in low LoA in optical imaging sensing applications and an open challenge in advanced 3D metrology tasks (CATALUCCI et al. 2022; ZHANG et al. 2021). Especially within large-scale metrology applications, camera exposure optimization is even more challenging when an inhomogeneous image intensity can be expected. For instance, this situation could arise in the acquisition of materials with a significant degree of specular reflection (see Subsection 2.4.1) and a complex geometry, such as car doors. Fig. 3.3 shows such an exemplary case and demonstrates that a single exposure time would not be sufficient to capture all features of a car door at once. It also shows that the spatial relationship between the features and the sensor must be considered to achieve an adequate image exposure.

To achieve an appropriate image exposure, a range of factors that affect scene illumination must be taken into account, such as focal length, working distance, and projector light intensity (see Subsection 2.4.1). This multi-dimensional problem demands a comprehensive spatial understanding of the captured scene to ensure optimal exposure time selection. For this reason, an operator is still required to manually adjust the camera exposure in complex large-scale coordinate metrology applications. Motivated by the idea that an expert can effectively parameterize the camera exposure time where other approaches fail, the present research assumes that a solution that emulates the operator behavior

has the greatest potential to address RQ3. Therefore, as a first step, the following generic behavior of the operator can be assumed. (MAGAÑA et al. 2023c)

- Observation (image exposure): Given an initial exposure time, the operator observes the local image exposure of the features to be captured.
- Action (exposure time and spatial relations): Based on the observed image exposure, the operator adjusts the exposure time taking into account the spatial relationships between the sensor and the features. Then, the expert triggers the measurement.
- Assessment (point cloud): The operator evaluates the measurement based on the resulting point cloud. The previous steps are repeated if the required point cloud quality is not achieved around the features of interest.

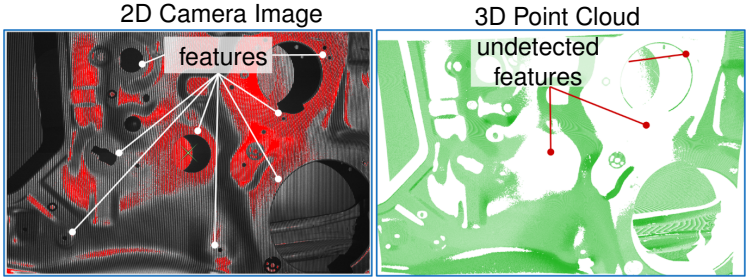


Figure 3.3: 2D camera image of a structured light sensor and the corresponding 3D point cloud of a car door measurement. The 2D images exemplify the qualitative illumination of the scene. The red pixels indicate an over saturation of the camera impeding the acquisition of the surface points around the features of interest. (modified figure from MAGAÑA et al. (2023c))

3.1.3.2 Solution Module Deliverable

The present thesis outlines the following deliverable that will allow to improve the LoA of automation of the sensor exposure time parameterization.

> A methodology that uses the sensor camera image exposure and the spatial relationships between sensor and surface to predict the expected measurement quality of a point cloud and estimate an appropriate exposure time.

3.2 Implementation – 1. Publication

PUB1	MAGAÑA et al. 2020a: „PoseNetwork: Pipeline for the Automated Generation of Synthetic Training Data and CNN for Object Detection, Segmentation, and Orientation Estimation“ MAGAÑA, WU, BAUER, and REINHART 2020

An end-to-end methodology for detecting and estimating the pose of an object of interest using a single camera image was introduced in PUB1. Fig. 3.4 depicts an overview of the proposed concept.

The publication first outlines a generic pipeline for automated dataset generation using a 3D CAD of the object of interest. The data generation pipeline consists of the following sub-steps:

- generation of synthetic images from multiple viewpoints
- image augmentation using variation of background and object position
- automated generation of annotations for image classification, segmentation, position, and orientation estimation

The designed pipeline demonstrated that the automated generation of such datasets can be achieved while ensuring their consistency using synthetic data.

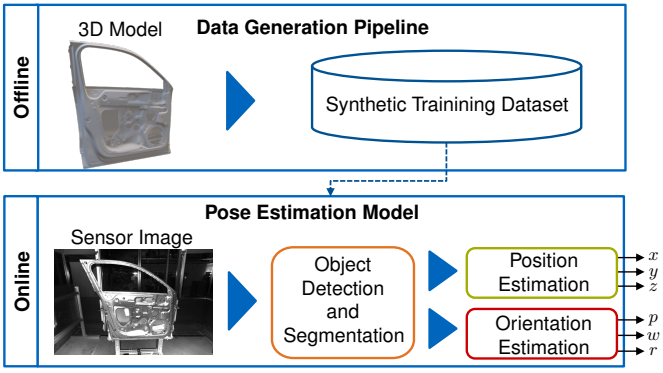


Figure 3.4: Concept proposed by MAGAÑA et al. (2020a) for training a pose estimation model based on a synthetic dataset generated using a 3D model.

Moreover, the publication proposes a multi-stage methodology comprising different image processing steps for estimating the pose of an object. The first module uses the CNN model proposed by HE et al. (2017) and REN et al. (2017) for detecting and segmenting objects. In a second step, MAGAÑA et al. (2020a)

introduce two independent sub-modules for estimating the position and orientation of the object based on the segmented image. The first sub-module object's employs the segmented image and applies the linear pinhole model for imaging cameras to approximate the object's position. The second module proposes a CNN architecture for estimating the object's orientation following a classification model. The architecture consists of three independent branches that allow the prediction of the orientation angle for each rotation axis. This architectural modularity allows the model to be trained and tuned individually for each axis and a defined orientation range.

The complete methodology was evaluated using an industrial ROMS for detecting a workpiece and estimating its relative position within its workspace. The evaluation results are more comprehensively discussed in Subsection 4.1.1. The scientific contributions of PUB1 are summarized as follows:

1. End-to-End methodology for 6D pose estimation and automated dataset generation using 3D solid models
2. Pipeline for automated synthetic dataset generation
3. Modular multi-stage methodology for object detection, segmentation, and pose estimation

The author's contribution to PUB1 is summarized in Tab. 3.1.

Table 3.1: Author's contribution to the publication MAGAÑA et al. (2020a)

Conceptual Design	Analysis and Evaluation	Manuscript Drafting	Data Collection and Software Development
70%	70%	80%	50%

3.3 Implementation – 2. Publication

PUB2	MAGAÑA et al. 2023a: „Viewpoint Generation Using Feature-Based Constrained Spaces for Robot Vision Systems“ MAGAÑA, DIRR, BAUER, and REINHART 2023

The publication of MAGAÑA et al. (2023a) addresses the VGP and demonstrates that this challenge can be handled as a purely geometric problem. First, the study introduces generic domain models for all core components of a ROMS to

promote the generalization of the outlined solutions. Then, the paper introduces a new mathematical formulation of the VGP based on $\mathcal{C}$-spaces. Based on these foundations, PUB2 shows that all viewpoint constraints from Subsection 2.3.2 can be analytically modeled and represented as 3D manifolds using different trigonometric approaches. For example, based on a geometric model of the sensor measurement volume, the pose of a feature, and a fixed sensor orientation, a core $\mathcal{C}$-space is synthesized based on linear algebra. Using this core $\mathcal{C}$-space, the rest of the viewpoint constraints are then characterized using various techniques such as linear algebra, trigonometry, geometric analysis, and *constructive solid geometry* (CSG) Boolean operations.

In addition, PUB2 shows that by using a consistent approach to model any viewpoint constraint multiple $\mathcal{C}$-spaces can be seamlessly combined. For example, the left image of Fig. 3.5 shows an occlusion-free $\mathcal{C}$-space (blue mesh) that integrates multiple constraints. The camera image and point clouds in Fig. 3.5 demonstrate that the feature and its entire geometry are within the measurement volume and can be captured despite the occluding object from a viewpoint with the $\mathcal{C}$-space. Furthermore, a real ROMS integrating two different sensors was used to validate the proposed viewpoint planning framework.

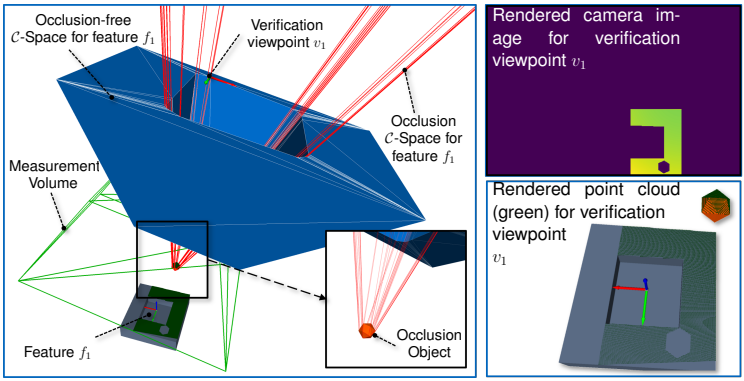


Figure 3.5: Left: Characterization of occlusion-free $\mathcal{C}$-space (blue mesh) for the acquisition of feature f_1. The occlusion $\mathcal{C}$-space (red wire-frame) represents the invalid solution space, which will prevent the valid acquisition of the whole feature. Right: Camera image and point cloud renderings for verifying viewpoint v_1 at the edge of the occlusion-free $\mathcal{C}$-spaces. (modified figure from MAGAÑA et al. (2023a))

This paper introduces the fundamental concepts of a generic framework that includes novel and efficient formulations to consistently characterize viewpoint constraints as $\mathcal{C}$-spaces. The key contributions of this scientific report are summarized as follows:

- Mathematical, model-based, and modular framework for formulating the VGP using $\mathcal{C}$-spaces and generic domain models

- Explicit formulation of up to nine viewpoint constraints in the Euclidean space using linear algebra, trigonometry, geometric analysis, and CSG Boolean operations
- Design of an academic reference scenario for benchmarking and further development

Tab. 3.2. summarizes the contribution of the thesis's author to PUB2.

Table 3.2: Author's contribution to the publication MAGAÑA et al. (2023a)

Conceptual Design	Analysis and Evaluation	Manuscript Drafting	Data Collection and Software Development
85%	85%	80%	90%

3.4 Implementation – 3. Publication

PUB3	MAGAÑA et al. 2023b: „Viewpoint Planning for Range Sensors Using Feature Cluster Constrained Spaces for Robot Vision Systems“ MAGAÑA, VLAEYEN, HAITJEMA, BAUER, SCHMUCKER, and REINHART 2023

While PUB2 deals with the computation of valid viewpoints for acquiring a single feature, PUB3 focuses on the second subquestion of RQ2 and introduces a holistic viewpoint planning strategy for the acquisition of multiple features, providing a solution for the SCP and the overall VPP.

MAGAÑA et al. (2023b) introduce two core modules to solve the VPP. In the first step, based on the foundations of $\mathcal{C}$-spaces the concept of ${}^{G}\mathcal{C}$-spaces is formally defined. Thus, a ${}^{G}\mathcal{C}$-spaces represents a solution space, where any viewpoint within it is valid for capturing a group of features while satisfying all individual viewpoint constraints. PUB3 shows that ${}^{G}\mathcal{C}$-spaces can be represented analogously as 3D manifolds and can be computed seamlessly by intersecting individual $\mathcal{C}$-spaces. Fig. 3.6 depicts the simplified characterization of two $\mathcal{C}$-spaces for capturing two groups of features. Moreover, the exemplary viewpoints at the edge of the respective ${}^{G}\mathcal{C}$-space demonstrate that the features including their geometry lie within the measurement volumes.

In the second step, a multi-stage viewpoint plan strategy aligned with the concept of ${}^{G}\mathcal{C}$-spaces is introduced to solve the SCP. The strategy considers a method for identifying potential groups of features that can be acquired together, a strategy for efficient characterization of ${}^{G}\mathcal{C}$-spaces, and a method for selecting viewpoints

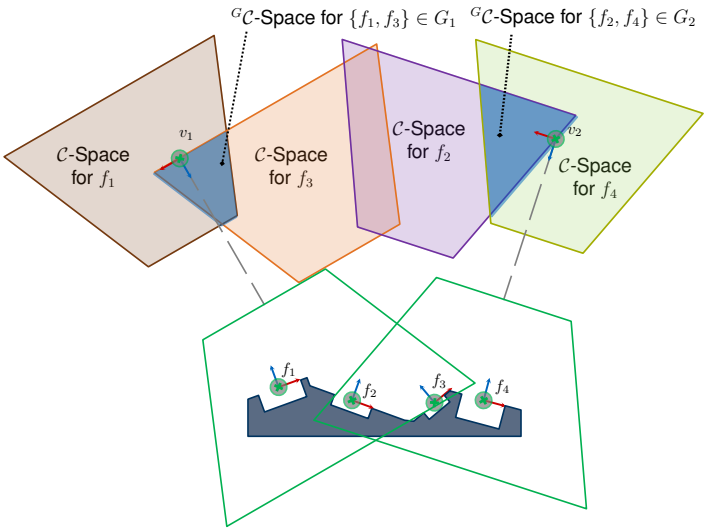


Figure 3.6: Simplified representation of two $^G\mathcal{C}$-spaces to acquire the group of features $\{f_1, f_3\} \in G_1$ and $\{f_2, f_4\} \in G_2$. The viewpoints v_1 and v_2 within the $^G\mathcal{C}$-spaces demonstrate that the respective features including their geometry lie within the measurement volume (modified figure from MAGAÑA et al. 2023b).

within the $^G\mathcal{C}$-spaces. The applicability and potential of the approach proposed are evaluated for automating dimensional metrology tasks using two different industrial measurement systems with different range sensors and manipulators. The results are discussed in Subsection 4.1.2.

The most relevant contributions of PUB3 are given as follows:

- Mathematical, generic and modularized formulation of the VPP to ease the transferability and promote the adaptability of the presented solutions with different ROMS and tasks.
- Characterization of $^G\mathcal{C}$-spaces as 3D manifolds based on $\mathcal{C}$-spaces, hence, inheriting some of their intrinsic advantages:
  - analytical, model-based, and closed-form solutions,
  - simple characterization based on CSG Boolean techniques,
  - infinite solutions for compensation of model uncertainties.
- Generic and modular viewpoint planning strategy that can be adapted to diverse vision tasks, ROMS, and viewpoint constraints.

The contribution to PUB3 by the author of this dissertation is in Tab. 3.3.

Table 3.3: Author's contribution to the publication MAGAÑA et al. (2023b)

Conceptual Design	Analysis and Evaluation	Manuscript Drafting	Data Collection and Software Development
90%	70%	70%	80%

3.5 Implementation – 4. Publication

PUB4	MAGAÑA et al. 2023c: „Exposure Time and Point Cloud Quality Prediction for Active 3D Imaging Sensors using Gaussian Processes“ MAGAÑA, SCHNEIDER, BENKER, ALTMANN, BAUER, and REINHART 2023

MAGAÑA et al. (2023c) address the challenge of estimating the exposure time and predicting the expected point cloud quality for active range sensor systems. To automate the parameterization of the sensor's exposure time, the following two submodules were conceptualized to replicate the expert approach.

1. Quality Estimation: The first module focuses on the design of a prediction model to estimate the local point cloud quality around a feature using the local image light intensity and the spatial relationships between the sensor and the features.

2. Exposure Time Estimation: Assuming that the expected point cloud quality around a feature can be estimated using the quality estimation module, this module's goal is to determine the appropriate exposure time by considering the nominal local image intensity and the spatial relationships between the sensor and the features.

In the context of PUB4 two data-based models were outlined to reproduce the required behavior of these submodules. Fig. 3.7 provides a graphical representation of the proposed models.

As a first step, generic model inputs and outputs were designed to model the camera image exposure, the point cloud quality, and the spatial system state. For example, the output variable p_{norm} was designed as a normalized metric to quantify the point cloud quality based on the number of points around a feature. The output metric takes into account feature geometry, acquisition depth, and camera resolution. Similarly the input variable i_{avg} was designed to characterize a local, feature-based image average intensity. In addition, positional and rotational inputs were introduced to model the spatial relationships between the sensor and the features.

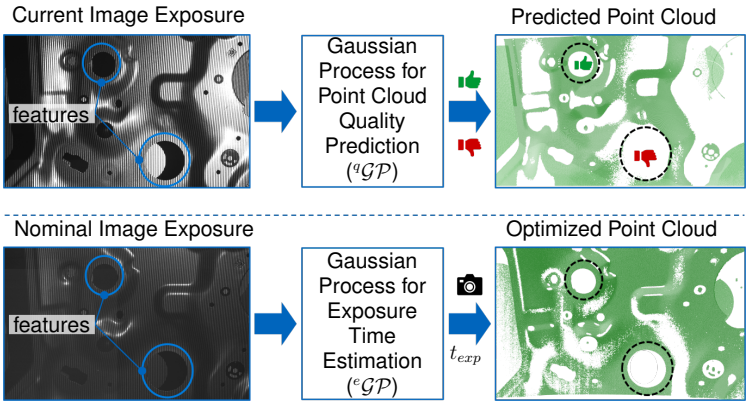


Figure 3.7: Methodology comprising two data-based models for predicting the measurement quality and for determining a valid exposure time for capturing two features. The upper workflow emulates the expert evaluation to determine the quality of point clouds around selected features. In contrast, the lower model depicts the action of an expert takes to select a proper exposure time based on the existing lighting conditions. (modified figure from MAGAÑA et al. (2023c))

In a second step, an incremental kernel design of the *Gaussian processes* (GPs) (see Subsection A.1.3.3) was followed to gradually increase the complexity of the models and ensure a proper selection of the required model inputs. The analysis systematically demonstrated that incorporating spatial relationships and nonlinear correlations significantly enhances the predictive performance of both models.

Finally, a GP regression model was outlined to predict the point cloud quality around a feature using the local camera image intensity and the 3D position between the sensor and a feature. Furthermore, a GP regression model was proposed to estimate the required exposure time for a given nominal image intensity and the 5D pose of a feature. The data for training and evaluation of both models were collected using an industrial *structured light sensor* (SLS) and a car door with different surface finishes.

The evaluation of the first model showed that the quality of the point cloud around a feature can be predicted with an RMSE[3] of 10 %. Moreover, the assessment of the second model showed that the exposure time could be predicted with an RMSE of 13 ms using an adjustment range of 1 ms to 350 ms. In addition, an exhaustive evaluation analyzed the performance of both models considering different scene lighting, a different active range sensor using stereo imaging, and different surface finishes. A more detailed discussion of the evaluation is presented in Subsection 4.1.3.

[3]The RMSE is a common metric used to evaluate regression models when predicting numerical values. The RMSE provides a quantification of how far the residuals, i.e., the difference between predicted and observed values, are on average from zero. (KUHN and JOHNSON 2013, p. 95)

The main contributions of the publication are summarized as follows:

- Synthesization of generic inputs and output variables for characterizing the point cloud quality, imaging, and spatial system state for active range image sensor systems.
- 4D GP generic kernel to predict the point cloud quality (number of points) around a feature given a local image intensity and its relative position (3D) to the sensor.
- 6D GP generic kernel to estimate the required exposure time for a feature's nominal local image intensity considering its relative position (3D) and rotation (2D) to the sensor.

Tab. 3.4. gives an overview of the author's contribution to PUB4.

Table 3.4: Author's contribution to the publication MAGAÑA et al. (2023c)

Conceptual Design	Analysis and Evaluation	Manuscript Drafting	Data Collection and Software Development
80%	70%	70%	70%

Chapter 4

Evaluation

This chapter summarizes in its first section the technical achievements of the proposed solution modules evaluated using an industrial *robot-based optical 3D measuring system* (ROMS) and a reference metrological task. Then, Section 4.2 presents an analysis of the economic viability of the overall approach.

4.1 Technical Evaluation

This section analyzes first the technical validity for each solution module in Subsections 4.1.1–4.1.3. Then, Subsection 4.1.4 provides an overall assessment of all solution modules in relation to the objective of this dissertation. The description of the technical experimental setup is given in A.3.1.

4.1.1 Workcell Calibration

The proposed methodology introduced by MAGAÑA et al. (2020a) was evaluated using the camera of the *structured light sensor* (SLS) to estimate the relative pose between the car door and the sensor using a single shot monochromatic image.

First, a 3D *computer-aided design* (CAD) model of the car door was used to automatically generate a dataset of 21,600 images. The synthetic dataset combined different randomization techniques such as different backgrounds, viewpoints, and scale factors. All *convolutional neural network* (CNN) models were trained exclusively with the generated synthetic dataset and using pre-trained models for image classification.

The performance and accuracy of the proposed models were assessed based on a comprehensive design of experiments considering 65 different camera images taken from different sensor poses and different rotations of the car door. The final evaluation showed that the object and segmentation models were able to detect the real car door in all cases and estimate its position with a standard deviation of 25 mm and up to 3.8° for the orientation. An overview of two exemplary input

images and output results for the different stages of the methodology is shown in Fig 4.1.

Taking into account the requirements outlined in Section 3.1, an assessment of the technical implementation is conducted. The strategy for generating synthetic datasets is highly generalizable, as it can be used with any 3D mesh model of an object. Furthermore, the automated process for generating datasets demonstrated a significant reduction in the time and expertise required to produce a high-quality dataset within a few hours, thereby enhancing the usability of the proposed data-driven approach.

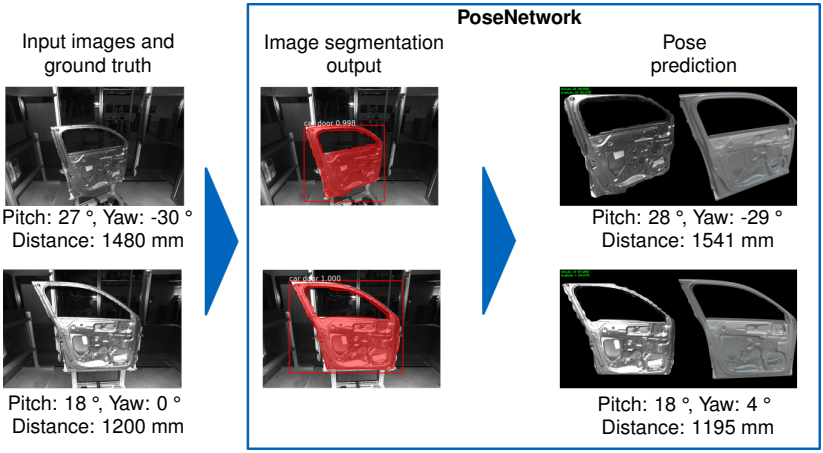


Figure 4.1: Evaluation overview: Prediction of a car door's orientation (yaw and pitch angles) and distance based on a grayscale image. (excerpt from MAGAÑA et al. 2020a)

4.1.1.1 Fulfillment of the Research Question and Solution Module Deliverable

The publication of MAGAÑA et al. (2020a) proposes a methodology for addressing RQ1 that allows an automated workcell alignment using single camera images obtained by the optical sensor of the ROMS. In addition, the proposed methodology addresses two challenging problems identified in Subsection 3.1.1 when executing an image-based alignment utilizing data-based models. These problems include an industrial-compliant approach for automatic dataset generation and an end-to-end methodology for pose estimation.

4.1.1.2 Limitations and Future Work

The proposed methodology showed potential for full automation of workcell alignment in the experimental setup of this thesis. However, future research should include a more comprehensive analysis of other objects, different scenes,

and benchmarking with other models. Moreover, further efforts should be dedicated to enhancing the dataset pipeline by integrating other image augmentation features such as lighting conditions, pixel augmentation (intensity, contrast, saturation), and image occlusion. The extension of such characteristics should increase the robustness and accuracy of the model predictions.

Moreover, the design of a holistic automatic calibration strategy was beyond the scope of this work. The full automation of this task requires the integration of additional modules. Future research should outline a path-planning strategy that allows the collision-free acquisition of multiple images in a safe space. In addition, a multi-stage calibration combining a coarse image-based calibration with a fine point cloud calibration would improve the overall accuracy of the approach (see Subsection 2.2.3). Finally, the models need to be extended to enable multi-object detection.

4.1.2 Viewpoint Planning

Based on the methods proposed by PUB2 and PUB3, a viewpoint planning framework was developed to generate automated viewpoints for the industrial ROMS presented in Subsection A.3.1. To evaluate the viewpoint planning framework, 15 different inspection tasks with up to 673 features were outlined in PUB3. The tasks included combinations of multiple features from the inside and outside of the door as well as different viewpoint constraints.

The viewpoint plans for the 15 inspection tasks were computed following the strategy outlined in PUB3. Fig. 4.2a shows the computed *feature-cluster constrained spaces* (${}^{G}\mathcal{C}$-spaces) for acquiring 50 circular features of the car door. To quantify the performance of the planning results, measurability and computational efficiency metrics were introduced. The measurability metric, which evaluates how many features could be captured, showed that the computed viewpoint plans achieved an average measurability index over 95 % for all tasks. Furthermore, the evaluation of the computational efficiency demonstrated that viewpoint plans requiring the acquisition of 673 features could be calculated within 5.3 min and with a measurability index over 97 %. The detailed results of each inspection scenario are presented in PUB3.

The evaluation of PUB2 and PUB3 demonstrated that spanning an infinite solution space is a powerful technique for explicitly solving the *Viewpoint Planning Problem* (VPP) within complex multi-feature scenarios involving modeling uncertainties. In addition, the framework proved to be hardware-agnostic as long as the domain models of manipulators, sensors, and workpieces can be modeled considering the assumptions and guidelines outlined in MAGAÑA et al. (2023a,b). For example, besides SLSs, the concept of *feature-based constrained spaces* ($\mathcal{C}$-spaces) and ${}^{G}\mathcal{C}$-spaces could be evaluated by different range sensors with varying principles of acquisition, a stereo sensor in PUB2 and a laser scanner in PUB3. Fig. 4.2b depicts the ${}^{G}\mathcal{C}$-spaces for capturing the surface of 3D cylinder features with a laser scanner.

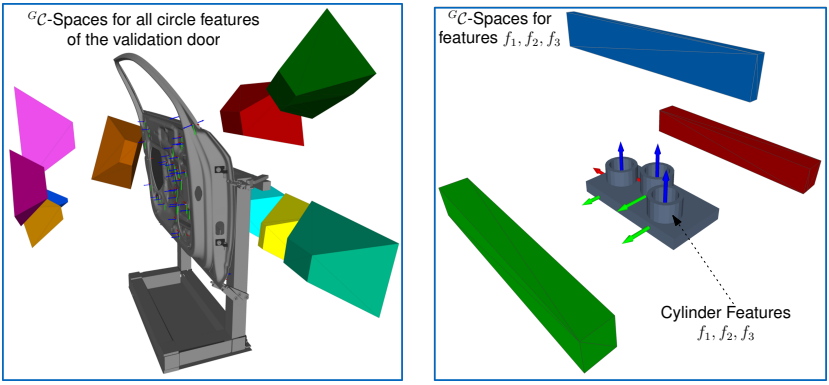


Figure 4.2: a) Representation of the ten characterized $^{G}\mathcal{C}$-spaces for acquiring 50 circle features of the car door using the SLS Zeiss Comet Pro AE. Approximate door and fixture dimensions in mm (length, width, height): $1200 \times 500 \times 1600$. b) Representation of three characterized $^{G}\mathcal{C}$-spaces for acquiring three 3D cylinder features using a laser scanner. Approximate probing object dimensions in mm (length, width, height): $120 \times 60 \times 34$. (modified from MAGAÑA et al. 2023b).

4.1.2.1 Fulfillment of the Research Question and Solution Module Deliverable

PUB2 and PUB3 followed a geometric approach to solve the VPP and introduced the fundamental components of a holistic viewpoint planning framework to tackle complex coordinate metrological tasks comprising multiple features with a ROMS. PUB2 focused on answering the first subquestion of RQ2 and investigated the influence of viewpoint constraints geometrically up to in 6D. Moreover, PUB2 outlined a modular framework for consistently characterizing all viewpoint constraints considered within the scope of the present thesis, see Subsection 2.3.2. Based on $\mathcal{C}$-spaces, and considering an adequate formulation of the overall VPP, PUB3 introduced a viewpoint planning strategy for computing a sufficient number of points that answers the second subquestion of RQ2.

4.1.2.2 Limitations and Future Work

The framework developed in this solution module effectively addressed the overall objective of this research. However, it is important to note that replicating this framework will require a significant effort for system integration tasks and comprehensive modeling of the domain components of ROMSs. For simpler tasks which require few constraints and features, this thesis recommends conducting a thorough analysis of the task to evaluate the use of alternative techniques. Examples of such techniques include those proposed by CHEN and LI (2004), GLORIEUX et al. (2020), KABA et al. (2017), and MITTAL and DAVIS (2007). The economic analysis of Section 4.2 provides an evaluation of the needed resources for replication, serving as a useful starting point to assess the proposed solution.

Moreover, the evaluation of PUB3 highlighted that some sensor poses were not able to capture all required features. Empirical measurements showed that the main cause of the failed measurements could be attributed to modeling uncertainties of the workcell kinematic model and manufacturing tolerances of the fixture. Some of the failed measurements could be corrected by selecting an alternative sensor pose within the corresponding ${}^{G}\mathcal{C}$-spaces. This characteristic embodies the intrinsic advantages of $\mathcal{C}$-spaces for compensating uncertainties without recomputing a new viewpoint plan. However, due to the technical limitations of the considered experimental setup, online adaptation could not be directly integrated into the implemented solution. For this reason, future implementations aiming at full automation should consider a technical module that allows the online adaptation of viewpoints.

In addition, future research should focus on the integration of additional constraints. For example, the characterization of overlapping constraints aligned to the concept of $\mathcal{C}$-spaces would help to ensure that an overlapping area between adjacent measurements can be satisfied. Integrating such constraints is particularly interesting for other vision tasks aiming at a full digitization and for techniques based on adjacent point cloud registration, such as BAUER et al. (2021b). In general, the results of the modeling approach should encourage the geometrical characterization of constraints that influence the planning of viewpoints. Such constraints may include, for example, sensor illumination, cycle times, or energy efficiency.

4.1.3 Sensor Parameterization

Selecting an adequate exposure time for an active range sensor system is a challenging task when an inhomogeneous light distribution caused by the light source and surface topology is present (see Fig. 3.3). To address this problem, PUB4 proposed a data-based approach using two *Gaussian process* (GP) regression models to predict the quality of a measurement and to predict the exposure time of a measurement given a nominal image intensity and spatial inputs. In this context, PUB4 considered the development and evaluation of the outlined models using the ROMS presented in Subsection A.3.1.

First a 1D GP regression model was trained to demonstrate the prediction performance that is obtained using only the local image intensity i_{avg}. The results of this model indicated that the point cloud quality can be estimated with an RMSE about 20 % and an R^2[1] of 78 %. By integrating spatial inputs that model the relative position between the sensor and the features, the RMSE of the point cloud quality could be improved up to 10 % and an R^2 of 93 %. A visualization of the model's mean prediction in the x-y plane, considering two light intensities and three different working distance positions, is depicted in

[1]The coefficient of determination R^2 is a statistical measure used in the evaluation of regression models to quantify the proportion of the information in the data that can be explained by the model using the model residuals as the basis for its calculation, i.e, the difference between observed and the prediction values. (KUHN and JOHNSON 2013, pp. 95–96; BACKHAUS et al. 2016, p. 72)

Fig. 4.3. These trends further confirm the motivation of this thesis, showing the nonlinear nature of the problem and the existence of local minima for the point cloud quality at different positions considering different image exposures. In addition, on the right of Fig. 4.4 the predictions of the quality model around different features is visualized.

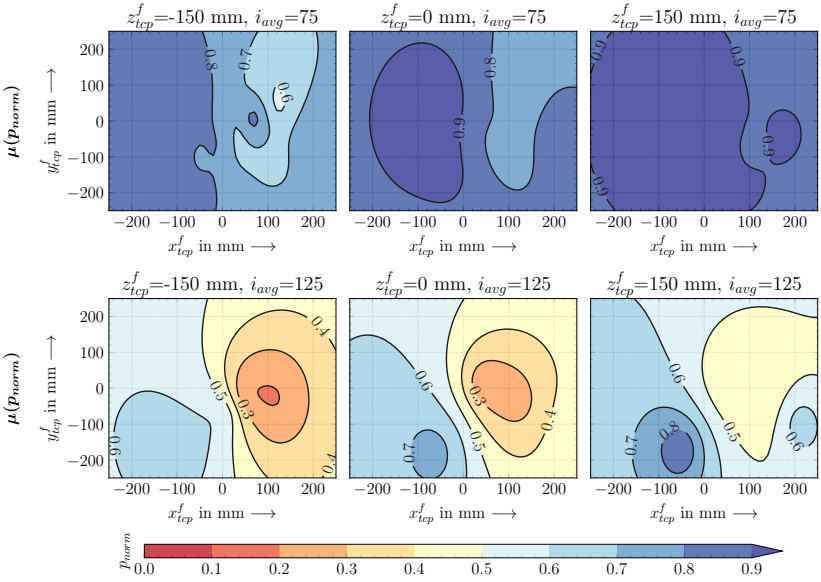


Figure 4.3: Discretized visualization of the mean prediction $\mu(p_{norm})$ in the x^f_{tcp} and y^f_{tcp} axes using the 4D GP quality regression model at three different depths $z^f_{tcp} = [-150, 0, 150]$ mm (left, middle, right) for two different light intensities $i_{avg} = [75, 125]$ (top, bottom) with a sensor orientation of $w^f_{tcp} = p^f_{tcp} = r^f_{tcp} = 0°$ for the SLS Zeiss Comet Pro AE. The graphs demonstrate that the point cloud density around a feature depends on the image intensity but also on the 3D relative position between a feature and the sensor. The trends confirm the nonlinear correlation between the point cloud quality, the image light intensity, and the spatial variables. (modified from MAGAÑA et al. 2023c)

Similarly, a 1D GP model was trained for predicting the exposure time using just the local image intensity around a feature. As expected, the prediction of this model showed to be ineffective. The estimated exposure time for achieving a desired image exposure had an RMSE about 75 ms and an R^2 of 7 % considering the setting range of the SLS 1 ms to 350 ms. By increasing the model complexity considering the relative position of the features to the sensor, the accuracy could be increased up to 33 ms. Finally, the best model performance was obtained by integrating positional and rotational inputs reaching an RMSE of 13 ms with an $R^2 = 97\,\%$. A graphical representation of the performance of the models is depicted in Fig. 4.4. The predicted exposure times are displayed on the left side, while the predicted point cloud quality values for different features are shown

on the right.

Both models were evaluated under different lighting conditions and car door surface finishes to assess their robustness. The findings showed that there was no significant variation in performance between both models when tested under different lighting configurations. These results confirmed the assumption of this thesis and previous findings, regarding the projector's intensity dominance over external light sources (EKSTRAND and ZHANG 2011; ZHANG 2020). Furthermore, the quality prediction model showed similar performance for features with different surface finishes. However, the exposure time prediction model showed a higher sensitivity to the surface finish, resulting in a poorer performance of the model with an RMSE of 21 ms. These results were expected since the model did not account for an input variable that characterizes the reflective properties of the door material.

In the context of generalization, the designed inputs/outputs and GP kernels, showed a high transferability, which proved to be equally effective for an alternative active range sensor using a stereo vision principle and a projector. Moreover, both models proved to be computationally efficient, given that the computation of 636 values corresponded to 0.6 s.

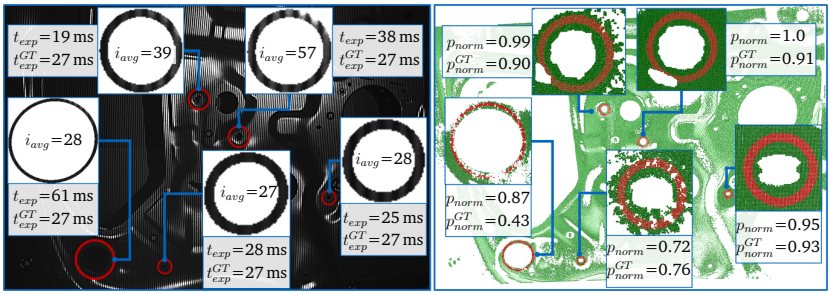


Figure 4.4: Performance overview of the GP models for five different features using the local average intensity (i_{avg}) and spatial relationships to the sensor. a) Comparison of the predicted exposure times t_{exp} using the GP for exposure time prediction for a exposure time with a ground truth (GT) of 27 ms. b) Comparison of the predicted point cloud density using the GP quality model with different GT values. (modified from MAGAÑA et al. 2023c)

4.1.3.1 Fulfillment of the Research Question and Solution Module Deliverable

PUB4 introduced a data-based approach consisting of two GP models to characterize the correlations between the camera exposure time, the point cloud density, the image light intensity, and the spatial relations of the acquisition scene. The models outlined in this solution module fulfill the defined deliverable of estimating the required exposure time to acquire dense point clouds around any arbitrary feature of a workpiece. Furthermore, the models of PUB4 addressed

RQ3 and demonstrate the presence of nonlinear correlations between the inputs and outputs of the models (see Fig. 4.3). The key findings showed that the spatial relationships between the sensor and the surface must be taken into account to guarantee a robust acquisition of the object surface and an adequate parameterization of the sensor exposure time.

4.1.3.2 Limitations and Future Work

When utilizing data-driven approaches, a large dataset is necessary to obtain an adequate performance of the trained models. The automation capabilities of an ROMS can considerably reduce the effort required for data collection. Moreover, PUB4 demonstrated that the GP models can be trained with a few spatially well-distributed observations (approximately 1500, as stated by MAGAÑA et al. 2023c). However, it is essential to consider the availability of time and expertise necessary for data collection and post-processing tasks when integrating data-driven methods. Therefore, future work should explore transfer learning techniques to reduce the effort of data collection and model retraining (TIGHINEANU et al. 2021).

The models underwent thorough evaluation and demonstrated potential for transferability to other features, diverse surface finishes, lighting conditions, and alternative active sensors. Nevertheless, it is crucial to acknowledge that their performance is initially restricted to the measurement setup considered in this thesis. For this reason, future research should consider a wider range of materials with different reflectance properties and diverse lighting conditions to investigate the robustness and validity of the models and the proposed inputs. To ensure replicability, the suggested models serve as a strong foundation. Nonetheless, it is crucial to carefully design the model and adapt it to the relevant system.

Finally, automating the selection of an appropriate exposure time based on a point cloud quality threshold requires a holistic strategy combining both models, which remains an open question beyond the scope of this work. This strategy should incorporate multi-objective optimization to obtain a global exposure time for acquiring multiple features simultaneously. Additionally, modeling the spatial correlations found within PUB4 as additional viewpoint constraints presents a promising challenge. By modeling the found spatial correlations aligned to the viewpoint planning framework of PUB2 and PUB3, measurement quality and exposure time constraints could be directly considered in the offline viewpoint planning.

4.1.4 Discussion

An integral technical evaluation regarding the overall improvement for programming ROMS to automate metrological tasks is assessed with this subsection.

The outlined approaches effectively addressed the identified challenges for each solution module. The obtained results demonstrated that the overall *level of*

automation (LoA) was increased to a *cognitive/information level of automation* (LoA_i) in the order of five, i.e., *supervision*. According to FROHM et al. (2008), this level demands a reduced intervention of the operator mostly in case of anomalies.

For instance, the methodology for image-based workcell alignment presented in PUB1 (see Subsection 4.1.1) showed to be a promising calibration strategy ordered at a $\text{LoA}_i = 5$. Using the outlined methods the workpiece could be identified using a camera image and its relative pose to the sensor could be estimated with a mean standard deviation up to 25 mm from a safe working distance of 1500 mm. These insights suggest that by extending the calibration strategy with the proposed enhancements (see Subsection 4.1.1.2), the robot could autonomously use the sensor's camera to estimate the pose of any artifact within the workcell.

In addition, the viewpoint planning framework (see Subsection 4.1.2) demonstrated that the programming effort for many inspection tasks could be significantly reduced to a level where no human intervention is required. However, the evaluation of PUB3 also showed that some inspection features could not be acquired for all planned viewpoints (see Subsection 4.1.2.2). In some cases the manual selection of alternative viewpoints within the $\mathcal{C}$-space was sufficient to find a valid solution. As the scope of this thesis was limited to the implementation of the solutions and not a complete system integration, the automated viewpoint selection was not integrated into the control software of the AIBox. Therefore, a holistic evaluation could not be performed. For this reason, the present work categorizes the overall achieved level of automation of the viewpoint planning framework as $\text{LoA}_i = 5$, which would require human attention to correct individual failed viewpoints.

For the last solution module, the evaluation of the data-based models within PUB4 showed a high accuracy with an R^2 above 90 % for predicting the point cloud quality of a measurement and estimating the corresponding required exposure time. These insights show that the proposed models have great potential to fully automate exposure time parameterization. However, the remaining model uncertainty shows that the proposed models and inputs do not fully characterize all the influencing factors that affect the exposure time setting and the point cloud quality. Thus, it is imaginable that a user will still be required to correct some outliers. For this reason, the cognitive LoA_i of the current solution is ordered also at a level of five.

The overall LoA of the ROMS could not be evaluated holistically due to the limited integration of the proposed solution modules. Nevertheless, the empirical evaluation of the solution modules showed their potential to automate the programming of the individual subtasks. Based on these insights and the results obtained from all publications, the overall cognitive LoA of ROMSs for automating coordinate metrology tasks was substantially improved by minimizing the human intervention up to a supervision level of five. Furthermore, addressing the outstanding concerns for each solution module could increase the LoA_i to seven,

indicating that the system could solve these metrological tasks autonomously.

4.2 Economic Evaluation

This section presents an economic analysis to demonstrate the benefits of the dissertation. First, Subsection 4.2.1 introduces an analysis for calculating the programming saving time-effort to automate metrology tasks. Based on these findings, Subsection 4.2.2 outlines the monetary benefits of the solutions taking into account economic methods.

4.2.1 Programming Saving Effort

This section introduces an exemplary calculation of the programming effort saved for each solution module using the outlined example from Subsection 1.2.2.

First, based on the evaluation results of PUB1, the workpiece alignment using a single camera image required an average of 15 s for each prediction. Therefore, an approximated time effort of

$$T_{align} \approx 1.5\,\text{min}$$

is considered realistic for calibrating a workpiece. The estimated time effort includes moving the robot to a safe position (60 s), capturing a camera image (15 s), and estimating the workpiece pose (15 s). Moreover, the computational efficiency of the viewpoint planning framework was evaluated in PUB3. As a result, the computation time of a viewpoint plan for the reference metrology task comprising 636 features corresponded to

$$T_{vp} = 5.3\,\text{min}.$$

Finally, for the sensor optimization solution module, it is assumed that each GP model must be run at least one time for each feature. The computation time for 636 predictions was estimated at 0.6 s for each model. In addition, the processing time for extracting the required inputs was estimated at 0.4 s for each feature. Therefore, the total time required to compute a valid exposure time for all measurements corresponds to

$$T_{par} = 2 \cdot 0.6\,\text{s} + 636 \cdot 0.4\,\text{s} = 256\,\text{s} = 4.3\,\text{min}.$$

Based on this analysis, the total computational effort for all solution modules corresponds to the sum of all calculated times

$$T_{aut} = T_{align} + T_{vp} + T_{par} \approx 11\,\text{min}.$$

A direct comparison with alternative approaches is not possible because the viewpoint planning problem lacks the design of benchmark scenarios. However, it is worth noting that the programming effort required for the approach proposed in this thesis can be compared to the empirical study conducted by BAUER et al. (2021a). In their paper, an online survey was completed by 27 experts who estimated an average time of 13.6 h to perform a similar inspection task using a ROMS as the one evaluated in this dissertation. The total estimated time took into account all activities related to the execution of the task, i.e., preparation, programming, measurement execution, data evaluation and post-processing. Since BAUER et al. (2021a) did not offer a detailed assessment of the individual processing times, this thesis conducted a separate analysis of their survey to breakdown the average time required for the programming of the ROMS. The following total programming effort was estimated:

$$T_{ref} = 450\,\text{min} \approx (7.5\,\text{h}). \tag{4.1}$$

This programming time does not include an explicit separation between online and offline programming. Moreover, it is consistent with the empirical experiments conducted in this thesis. Therefore, the effort estimation derived from BAUER et al. (2021a) is considered a valid reference metric within this dissertation.

Taking into account that the the full automation of the tasks of was not achieved for either the viewpoint planning (95 % planning effectiveness) or the sensor parameterization (90 % model fitness), it is assumed that a minimal human intervention may still be necessary to correct up to 10% of the planned viewpoints and predicted exposure times. The total programming effort can then be corrected using the reference time from Eq. 4.1 and assuming that the time effort is proportional to the number of features. Therefore, a more realistic computation is given for a semi-automatized programming process as follows:

$$T_{semi-aut} = T_{aut} + 10\,\% \cdot T_{ref} \approx 56\,\text{min}.$$

Although this effort is five times larger than the first estimation, it still represents a considerably programming saving factor given as follows:

$$p_{saving} = 1 - \frac{T_{semi-aut}}{T_{ref}} = 87.6\,\%. \tag{4.2}$$

In addition, it must be noted that the estimated computation times in this subsection provide a rough estimate of the programming time required to solve the individual problems. For example, the individual estimation of the workcell calibration effort was not considered in the estimation of the reference programming time t_{ref}. According to empirical experiments, the programming time for workpiece calibration corresponds to 10 min to 15 min. Therefore, a higher saving programming factor should be expected when considering this effort in the calculation of p_{saving}.

Furthermore, this analysis assumes that all solution modules have been fully integrated into the proprietary software of the ROMS and that the corresponding models have been previously implemented, trained, and parameterized.

4.2.2 Discussion

The most notorious economic benefit of the results of this thesis is given implicitly by the reduction of the programming effort. However, to outline the real economic effort to develop and integrate the outlined solution modules, a comprehensive analysis is presented in Subsection A.3.2. First, a detailed breakdown of the investment costs (A.3.2.1) showed that nearly 80 % of the costs must be allocated to software development for implementing the solution modules. Then, based on the calculated programming savings from Subsection 4.2.1, a payback period of approximately two years can be expected considering 96 measurement tasks (see Subsection A.3.2.2) or after the inspection of 118,228 features (see Subsection A.3.2.3).

Although these numbers may seem high for end-users who require a large number of probes, it is important to note that they are typical in quality inspection departments of automotive manufacturers. Therefore, for end-users with a high number of measuring tasks, implementing the proposed modules can be profitable. Furthermore, a system integrator can also capitalize on the economic advantages by developing a solution based on the outlined modules and providing a comprehensive end-to-end solution to their clients.

The economic assessment of this section has shown that improving the programming of ROMS using the proposed solutions is not only technically beneficial, but also profitable. It is important to emphasize that the economic analysis presented in this subsection is intended to provide a rough estimate of the magnitude of the costs associated with the proposed approach. Therefore, for a more accurate estimate of the economic benefits, a more comprehensive evaluation should consider dynamic methods with variable costs per year and qualitative and quantitative metrics, e.g., increase in productivity, utilization of the framework with other machines, quality assurance, relief of employees from repetitive tasks. If such effects can be accurately quantified, the cost-benefit ratio for improving the programming of ROMS is expected to be achieved early.

4.3 Overall Discussion and Reflection

Programming valid measurement routines for ROMSs is a complex task that still relies on manual intervention. In Section 1.2, the most challenging tasks that require expert input, including workcell calibration, viewpoint planning, and sensor parameterization. These tasks are complex and repetitive, requiring skilled personnel. Automating these tasks requires a thorough understanding of the system relationships and external factors, such as lighting conditions and design uncertainties. It also requires the development of technical solutions that end-users can apply in real-world scenarios. This thesis addresses each of these identified solution modules individually by formalizing the implicit and explicit knowledge required to automate them.

Calibrating workcells for robot-based systems is a crucial task for any production system. It often requires manual intervention to prevent collisions, which is why skilled operators are typically tasked with the job. A vision system capable of detecting objects is necessary for this task. The latest research, discussed in Subsection 2.2, shows that solutions based on *artificial neural networks* (ANNs) can detect these objects. However, an end-to-end solution that can detect and estimate the pose of an object is needed. For this reason, this thesis proposed a strategy comprising multiple ANNs to estimate the 6D pose of an object based on a sensor image. Moreover, this research revealed that ANNs can be trained effectively using synthetic datasets only, without the need for collecting and labeling large amounts of real-world data.

The second module, which focused on viewpoint planning, aimed to address the challenge of modeling the solution space and constraints for positioning sensors. While current state-of-the-art solutions (as discussed in Subsection 2.3) rely on heuristic algorithms that perform well, they have limitations when dealing with uncertainties. Therefore, this thesis proposed to use trigonometric and geometric methods to model all viewpoint constraints, both explicit and implicit, in order to formalize expert knowledge. By explicitly modeling all constraints, the solution space of a viewpoint could be modeled and uncertainties could be considered. This approach demonstrated that viewpoint planning can be solved effectively on the basis of solution spaces and that the task can be fully automated.

The last module dealt with the parameterization of sensors, which has been a persistent issue when it comes to utilizing optical systems. Specifically, projectors used in active systems have shown that spatial properties, material reflection properties, and projector intensity, among other factors, must be taken into account. Additionally, the non-linear relationships involved in this problem make it challenging to model analytically. Subsection 2.4 summarized different solutions available, but these require user intervention or system modifications, which can be impractical in a production context. To address this issue, this thesis introduced data-driven models that were trained using past measurements, as well as the lighting conditions and spatial relationships of the measurement setup. The results of this research showed that these models were able to formalize user skills by taking into account spatial and imaging understanding of the setup, which enabled an effective and automated parameterization process.

The presented solutions have demonstrated the possibility of formalizing missing knowledge, leading to the automation of individual tasks. This automation enables experts to focus on more creative endeavors, without the burden of programming repetitive tasks for ROMSs. Although these solutions address the challenges associated with automating programming, it is important to note that significant effort is required to implement and configure the necessary modules. However, once this initial investment is made, the solutions are both technically advantageous and economically feasible.

Chapter 5

Conclusion

5.1 Summary

Over the past three decades, robot-based systems have emerged as a key technology for tackling some of the challenges current production confronts, such as customized manufacturing, shorter cycle times, and flexibility. Especially in the field of quality assurance, *robot-based optical 3D measuring systems* (ROMSs), composed of an industrial robot and an optical range sensor, have the potential to automate complex inspection processes requiring vision capabilities (e.g., dimensional metrology, surface reconstruction, surface quality inspection) and enhance the productivity of these tasks. However, programming ROMS is still considered a challenging and resource-intensive task demanding extensive domain knowledge from robotics, optics, metrology, and computer vision. For these reasons, this thesis addressed some of the most pressing concerns regarding the automation of ROMS. Within four publications three independent solution modules were outlined to enhance the *level of automation* (LoA) of the system. This was achieved by improving the programming of the following subtasks: workcell calibration, viewpoint planning, and sensor parameterization.

The workcell calibration deals with the identification and pose estimation of any relevant object within the robot workspace that may affect the programming of the ROMS. For this subtask, a multi-stage strategy was outlined in PUB1 that integrates multiple *convolutional neural network* (CNN) models to identify and estimate the pose of an object based on a single camera image captured by the sensor of the ROMS. Additionally, a pipeline was proposed to automate this process by using synthetic data generated from CAD models, in order to mitigate the burdens of generating a training dataset. The CNN models were fully trained on synthetic datasets, demonstrating the potential to fully automate this laborious task while maintaining a good performance. The evaluation showed that the proposed strategy was able to autonomously identify a workpiece within the robot's workspace and estimate its pose with an accuracy of 25 mm. These results show the potential of such a strategy to eliminate any manual intervention during the workcell calibration process and increase the LoA of this task.

The second module dealt with the computation of a sufficient number of viewpoints to capture a set of inspection features of an object, taking into account various system constraints, such as sensor imaging parameters, occlusion, feature geometry, robot workspace, and model uncertainties. Due to its complexity, this task is often programmed manually by an expert. To address this challenge and increase the LoA of the viewpoint planning task, the present research suggest modularizing the problem into two stages. The first step, as detailed in PUB2, involves the analytical formulation of the overall problem and the use of solution spaces to characterize viewpoint constraints. PUB2 showed that nine constraints can be explicitly modeled using geometric and trigonometric relations. The second subproblem is addressed by PUB3, which introduces a strategy for computing an adequate number of viewpoints to perform an inspection task based on the proposed geometric modeling approach. A comprehensive evaluation of the viewpoint planning framework showed that valid viewpoints could be automatically found for different inspection tasks with different constraints, range sensor types, and workpieces. The findings of PUB2 and PUB3 indicate a significant increase in the LoA of this subtask and a high degree of generalizability for the framework developed.

The last solution module focused on the automatic parameterization of the camera's exposure time for active sensors to ensure the successful acquisition of point clouds with highly reflective surfaces. Sensor parameterization is complex nonlinear problem, which is particularly challenging due to the inhomogeneous illumination of the target surface caused by factors such as the asymmetric configuration of the light source, sensor position, and workpiece surface topology. To address this task, PUB4 outlined two *machine learning* (ML) models. The first model predicts the quality of measurements, while the second model estimates a valid exposure time. The modeling and evaluation of the models demonstrated that to properly estimate the quality and exposure time of the point cloud, it is essential to model to consider the spatial relationships between the sensor and features. Furthermore, the models demonstrated significant transferability between various surface finishes, illumination conditions, and active range sensors. The findings from PUB4 illustrate that an appropriate exposure time can be determined automatically, effectively decreasing the necessary manual labor for this task.

In conclusion, Chapter 4 presented an integral technical and economic assessment of all solution modules. The outcomes of the evaluation demonstrate a considerable improvement in the overall programming automation of ROMS. The analysis indicates that the proposed modules could enhance the system's productivity by lowering the programming effort by up to 88 % compared to a manual reference scenario. Additionally, a conservative cost-benefit analysis shows that the current solution modules are economically viable once the programming of viewpoint plans to measure the equivalent of 237 doors with 500 features is taken into account.

5.2 Outlook

The limitations and technical aspects of the individual solution modules were discussed in Subsections 4.1.1.2, 4.1.2.2, and 4.1.3.2. In the context of a more holistic evaluation, the full integration of all solution modules was beyond the scope of this thesis. The orchestration of all modules remains an open issue and a prerequisite for achieving the complete autonomy of ROMSs. Towards this goal, this thesis considers three research directions that remain to be addressed: hardware-independent integration, knowledge management, and validation of autonomous programs within industrial applications.

The use of hardware agnostic solutions is a common problem in the development of industrial automation software, which has motivated associations composed of end users and system vendors to work on open communication protocols, domain models, and exchange formats for modeling production processes and field devices, e.g., Asset Administration Shell, OPC-UA Foundation, AutomationML. Although some of these information models and communication technologies have gained acceptance in individual industries and regions, standards for the production and metrology have not been established in either industry or academia (SCHMITT et al. 2016; MÜLLER and KUTZBACH 2019, p. 521). Additionally, current research shows that concepts for modeling and programming field devices and manufacturing processes to increase the flexibility and productivity of production systems are still the focus of ongoing investigations (BAUER et al. 2019a; HAMMERSTINGL 2020; HEUSS et al. 2022). Therefore, future work aiming at using or enhancing the present solution modules should review the current research and standardization of communication protocols and domain models. This is necessary to develop a software framework that can effortlessly integrate into hardware-agnostic production. A first step towards the unified integration of all modules and knowledge management for ROMSs was investigated in MAGAÑA et al. (2020b). The publication proposed a knowledge base to centralize and formalize domain knowledge at different levels of abstraction for programming ROMSs. Although the results of MAGAÑA et al. (2020b) provided the basis for a hardware-agnostic implementation of some modules of this thesis, the domain models proposed by the authors still need to be aligned with a standard or industry specification.

On an organizational level, another pressing issue related to the automatic creation of programs generated by *artificial intelligence* (AI) is their evaluation considering different aspects such as safety, robustness, traceability, among others (HOFFMANN et al. 2021). Although AI driven solutions have shown great potential to solve some of the open problems addressed within this thesis, it should be recalled that some of the algorithms outlined rely on heuristic and stochastic methods. Therefore, there is a possibility that the results of the applied methods may not be deterministic, and their traceability can be compromised. Hence, further research must delve into how the suggested AI-based solution modules can be assessed or expanded to assure an easy application in industrial settings.

Furthermore, enhancing the LoA of production systems would ultimately aid in tackling some of the worldwide challenges that are impacting manufacturing, including demographic shifts, reshoring production, and constantly evolving product life cycles. However, the short-term impact on critical global trends remains to be analyzed. To this end, the goals set by the Agenda for Sustainable Development adopted by the United Nations (UNITED NATIONS 2022) could be used to guide the development of more sustainable and resource-efficient production systems. For example, in the context of this thesis, the automation of the sensor parameterization would significantly reduce the use of the sensor and improve its energy consumption. Furthermore, the viewpoint planning could be extended to integrate energy constraints and optimize the robot path plan. Therefore, a well-founded analysis to measure the system's energy consumption remains to be investigated considering appropriate methodologies, for example LIEBL (2020). Based on such an analysis, the benefits of automation could be transparently outlined, and concrete actions to improve the energy efficiency of the ROMS could be further derived.

Finally, the results of this work should stimulate future work to evaluate the transferability of the solution modules to similar non- or industrial tasks, such as vision-based robot self-calibration (LUX-GRUENBERG 2020), bin-picking applications (DIRR et al. 2023), robot-assisted surgery (SU et al. 2021), navigation (DUFEK et al. 2021), among others. The solution modules proposed in this thesis provide a stepping stone for researchers and decision-makers to automate similar tasks and be a helpful aid in evaluating their cost-benefit ratio.

Bibliography

ABELE and REINHART 2011
Abele, E.; Reinhart, G.: *Zukunft der Produktion. Herausforderungen, Forschungsfelder, Chancen*. München: Hanser. 2011. ISBN: 9783446425958.

ABRAMS and ALLEN 1992
Abrams, S.; Allen, P. K.: Sensor planning in an active robotic work cell. In: *Sensor Fusion IV: Control Paradigms and Data Structures*. Ed. by Schenker, P. S. SPIE Proceedings. SPIE. 1992, pp. 267–277. DOI: 10.1117/12.57928.

ABRAMS et al. 1999
Abrams, S.; Allen, P. K.; Tarabanis, K.: Computing Camera Viewpoints in an Active Robot Work Cell. *The International Journal of Robotics Research* 18 (1999) 3, pp. 267–285. DOI: 10.1177/02783649922066204.

BABAIE et al. 2015
Babaie, G.; Abolbashari, M.; Farahi, F.: Dynamics range enhancement in digital fringe projection technique. *Precision Engineering* 39 (2015), pp. 243–251. ISSN: 01416359. DOI: 10.1016/j.precisioneng.2014.06.007.

BACKHAUS 2016
Backhaus, J. C. S.: Adaptierbares aufgabenorientiertes Programmiersystem für Montagesysteme. PhD thesis. Technical University of Munich. 2016.

BACKHAUS et al. 2016
Backhaus, K.; Erichson, B.; Plinke, W.; Weiber, R.: *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. 14. ed. 2016. Berlin, Heidelberg: Springer Berlin Heidelberg and Imprint: Springer Gabler. 2016. ISBN: 978-3-662-46076-4.

BASLER 2021
Basler, D.: *Neuronale Netze Mit C# Programmieren. Mit Praktischen Beispielen Fur Machine Learning Im Unternehmenseinsatz*. Cincinnati: Hanser. 2021. ISBN: 978-3-446-46426-1.

BATCHELOR 2012
Batchelor, B. G.: *Machine Vision Handbook*. London: Springer London. 2012. ISBN: 978-1-84996-168-4. DOI: 10.1007/978-1-84996-169-1.

BAUER and MAGAÑA 2019
Patent: DE 10 2019 114 572 B3 pat. (01.10.2020). 2019. Bauer, P.; Magaña, A.: „Verfahren und System zur dreidimensionalen Erfassung einer Szene“.

BAUER et al. 2019a
Bauer, P.; Gonnermann, C.; Magaña, A.; Reinhart, G.: Autonome Prüfsysteme in der digitalen Fabrik: Skill-basierte Modellierung in der geometrischen Qualitätsprüfung für roboterbasierte Messsysteme. *wt Werkstattstechnik online* 109 (2019) 5, pp. 321–328.

BAUER et al. 2019b
Bauer, P.; Magaña, A.; Reinhart, G.: Free-Form Surface Analysis and Linking Strategies for High Registration Accuracy in Quality Assurance Applications. *Procedia CIRP* 81 (2019), pp. 968–973. ISSN: 22128271. DOI: 10.1016/j.procir.2019.03.236.

BAUER et al. 2020a
Bauer, P.; Li, F.; Magaña Flores, A.; Reinhart, G.: Accuracy Analysis of Alignment Methods based on Reference Features for Robot-Based Optical Inspection Systems. *Procedia CIRP* 93 (2020), pp. 1115–1120. ISSN: 22128271. DOI: 10.1016/j.procir.2020.04.105.

BAUER et al. 2020b
Bauer, P.; Fink, F.; Magaña, A.; Reinhart, G.: Spatial interactive projections in robot-based inspection systems. *The International Journal of Advanced Manufacturing Technology* 107 (2020) 5-6, pp. 2889–2900. ISSN: 0268-3768. DOI: 10.1007/s00170-020-05220-1.

BAUER et al. 2021a
Bauer, P.; Gottschall, L.; Magaña, A.; Hofer, A.; Reinhart, G.: A Technological and Economic Potential Analysis of Measurement Systems in Geometrical Quality Assurance. *Procedia CIRP* 104 (2021), pp. 1263–1268. ISSN: 22128271. DOI: 10.1016/j.procir.2021.11.212.

BAUER et al. 2021b
Bauer, P.; Heckler, L.; Worack, M.; Magaña, A.; Reinhart, G.: Registration strategy of point clouds based on region-specific projections and virtual structures for robot-based inspection systems. *Measurement* 185 (2021), p. 109963. DOI: 10.1016/j.measurement.2021.109963.

BAUER et al. 2022
Bauer, P.; Schmitt, S.; Dirr, J.; Magaña, A.; Reinhart, G.: Intelligent predetection of projected reference markers for robot-based inspection systems. *Production Engineering* 16 (2022) 5, pp. 719–734. ISSN: 0944-6524. DOI: 10.1007/s11740-022-01118-x.

BAUER and WARSCHAT 2021
Bauer, W.; Warschat, J.: *Smart Innovation durch Natural Language Processing.*

Mit Künstlicher Intelligenz die Wettbewerbsfähigkeit verbessern. München: Hanser. 2021. ISBN: 978-3-446-46606-7.

BEASLEY and CHU 1996
Beasley, J.; Chu, P.: A genetic algorithm for the set covering problem. *European Journal of Operational Research* 94 (1996) 2, pp. 392–404. ISSN: 03772217. DOI: 10.1016/0377-2217(95)00159-X.

BESL 1988
Besl, P. J.: Active, optical range imaging sensors. *Machine Vision and Applications* 1 (1988) 2, pp. 127–152. ISSN: 0932-8092. DOI: 10.1007/BF01212277.

BESL and MCKAY 1992
Besl, P. J.; McKay, N. D.: Method for registration of 3-D shapes. In: *Sensor fusion IV: control paradigms and data structures*. Vol. 1611. 1992, pp. 586–606.

BEYERER et al. 2016
Beyerer, J.; Puente León, F.; Frese, C.: *Machine Vision*. Berlin, Heidelberg: Springer Berlin Heidelberg. 2016. ISBN: 978-3-662-47793-9. DOI: 10.1007/978-3-662-47794-6.

BISHOP 2006
Bishop, C. M.: *Pattern recognition and machine learning*. Information science and statistics. New York: Springer. 2006. ISBN: 0387310738.

BLESSING and CHAKRABARTI 2009
Blessing, L. T. M.; Chakrabarti, A.: *DRM, a design research methodology*. Dordrecht and London: Springer. 2009. ISBN: 978-1-84882-586-4.

BOLYA et al. 2019
Bolya, D.; Zhou, C.; Xiao, F.; Lee, Y. J.: YOLACT: Real-Time Instance Segmentation. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE. 2019, pp. 9156–9165. ISBN: 978-1-7281-4803-8. DOI: 10.1109/ICCV.2019.00925.

BRACHMANN et al. 2016
Brachmann, E.; Michel, F.; Krull, A.; Ying Yang, M.; Gumhold, S., et al.: Uncertainty-driven 6d pose estimation of objects and scenes from a single rgb image. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 3364–3372.

BRADSKI 2000
Bradski, G.: The OpenCV Library. URL: https://github.com/opencv/opencv (visited on 12/03/2023).

BRAUER 2014
Brauer, J.: Human Pose Estimation with Implicit Shape Models. PhD thesis. Karlsruhe Institut für Technologie. 2014. DOI: 10.13140/RG.2.1.4074.6961.

BRECHER and WECK 2021
Brecher, C.; Weck, M.: *Werkzeugmaschinen Fertigungssysteme*. 9. Auflage. VDI-Buch. Berlin: Springer Vieweg. 2021. ISBN: 978-3-662-46568-4.

CATALUCCI et al. 2022
Catalucci, S.; Thompson, A.; Piano, S.; Branson, D. T.; Leach, R.: Optical metrology for digital manufacturing: a review. *The International Journal of Advanced Manufacturing Technology* 120 (2022) 7-8, pp. 4271–4290. ISSN: 0268-3768. DOI: 10.1007/s00170-022-09084-5.

CHEN and LI 2004
Chen, S. Y.; Li, Y. F.: Automatic sensor placement for model-based robot vision. *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society* 34 (2004) 1, pp. 393–408. ISSN: 1083-4419. DOI: 10.1109/TSMCB.2003.817031.

CHEN et al. 2011
Chen, S.; Li, Y.; Kwok, N. M.: Active vision in robotic systems: A survey of recent developments. *The International Journal of Robotics Research* 30 (2011) 11, pp. 1343–1377. DOI: 10.1177/0278364911410755.

CHOLLET 2018
Chollet, F.: *Deep learning with Python*. Shelter Island New York: Manning Publications Co. 2018. ISBN: 9781617294433.

CHOLLET et al. 2015
Chollet, F. et al.: Keras. URL: https://github.com/fchollet/keras (visited on 12/03/2023).

COWAN and BERGMAN 1989
Cowan, C. K.; Bergman, A.: Determining the camera and light source location for a visual task. In: *Proceedings, 1989 International Conference on Robotics and Automation*. IEEE Comput. Soc. Press. 1989, pp. 509–514. ISBN: 0-8186-1938-4. DOI: 10.1109/ROBOT.1989.100037.

COWAN and MODAYUR 1993
Cowan, C. K.; Modayur, B.: Edge-based placement of camera and light source for object recognition and location. In: *[1993] Proceedings IEEE International Conference on Robotics and Automation*. IEEE Comput. Soc. Press. 1993, pp. 586–591. ISBN: 0-8186-3450-2. DOI: 10.1109/ROBOT.1993.291895.

COWAN and KOVESI 1988
Cowan, C. K.; Kovesi, P. D.: Automatic sensor placement from vision task requirements. *IEEE Transactions on Pattern Analysis and machine intelligence* 10 (1988) 3, pp. 407–416. ISSN: 01628828. DOI: 10.1109/34.3905.

DAWSON-HAGGERTY ET AL. 2022
Dawson-Haggerty et al.: trimesh. URL: https://trimsh.org/ (visited on 12/03/2023).

DENG 2014
Deng, L.: Deep Learning: Methods and Applications. *Foundations and Trends® in Signal Processing* 7 (2014) 3-4, pp. 197–387. ISSN: 1932-8346. DOI: 10.1561/2000000039.

DERIGENT et al. 2006
Derigent, W.; Chapotot, E.; Ris, G.; Remy, S.; Bernard, A.: 3D Digitizing Strategy Planning Approach Based on a CAD Model. *Journal of Computing and Information Science in Engineering* 7 (2006) 1, pp. 10–19. ISSN: 1530-9827. DOI: 10.1115/1.2410023.

DIETZ 2019
Dietz, T.: Knowledge-based cost-benefit analysis of robotics for SME-like manufacturing. PhD thesis. University of Stuttgart. 2019.

DIN 55350 2021
DIN 55350: Begriffe zum Qualitätsmanagement. Berlin. DOI: 10.31030/3270647.

DIN EN ISO 10218-1 2012
DIN EN ISO 10218-1: *Robots and robotic devices - Safety requirements for industrial robots - Part 1: Robots*. Berlin: Beuth Verlag GmbH. 2012. DOI: 10.31030/1733801.

DIN EN ISO 10360-13 2023
DIN EN ISO 10360-13: *Geometrical product specifications (GPS) - Acceptance and reverification tests for coordinate measuring systems (CMS) - Part 13: Optical 3D CMS*. Berlin: Beuth Verlag GmbH. 2023. DOI: 10.31030/3161374.

DIN EN ISO 9000 2015
DIN EN ISO 9000: Quality management systems – Fundamentals and vocabulary. Berlin. DOI: 10.31030/2325650.

DIRR et al. 2023
Dirr, J.; Gebauer, D.; Yao, J.; Daub, R.: Automatic Image Generation Pipeline for Instance Segmentation of Deformable Linear Objects. *Sensors (Basel, Switzerland)* 23 (2023) 6. DOI: 10.3390/s23063013.

DU et al. 2021
Du, G.; Wang, K.; Lian, S.; Zhao, K.: Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review. *Artificial Intelligence Review* 54 (2021) 3, pp. 1677–1734. ISSN: 0269-2821. DOI: 10.1007/s10462-020-09888-5.

DUFEK et al. 2021
Dufek, J.; Xiao, X.; Murphy, R. R.: Best Viewpoints for External Robots or Sensors Assisting Other Robots. *IEEE Transactions on Human-Machine Systems* 51 (2021) 4, pp. 324–334. ISSN: 2168-2291. DOI: 10.1109/THMS.2021.3090765.

DUVENAUD 2014
Duvenaud, D. K.: Automatic Model Construction with Gaussian Processes. PhD thesis. University of Cambridge. 2014.

EKSTRAND and ZHANG 2011
Ekstrand, L.; Zhang, S.: Autoexposure for three-dimensional shape measurement using a digital-light-processing projector. *Optical Engineering* 50 (2011) 12, p. 123603. ISSN: 0091-3286. DOI: 10.1117/1.3662387.

ELLENRIEDER et al. 2005a
Ellenrieder, M. M.; Krüger, L.; Stößel, D.; Hanheide, M.: A versatile model-based visibility measure for geometric primitives. In: *Scandinavian Conference on Image Analysis*. 2005, pp. 669–678.

ELLENRIEDER et al. 2005b
Ellenrieder, M. M.; Wohler, C.; d'Angelo, P.: Reflectivity function based illumination and sensor planning for industrial inspection. In: *Optical Measurement Systems for Industrial Inspection IV*. Ed. by Osten, W.; Gorecki, C.; Novak, E. L. SPIE Proceedings. SPIE. 2005, p. 89. DOI: 10.1117/12.612547.

ERDEM and SCLAROFF 2006
Erdem, U. M.; Sclaroff, S.: Automated camera layout to satisfy task-specific and floor plan-specific coverage requirements. *Computer Vision and Image Understanding* 103 (2006) 3, pp. 156–169. DOI: 10.1016/j.cviu.2006.06.005.

FENG et al. 2014
Feng, S.; Zhang, Y.; Chen, Q.; Zuo, C.; Li, R.; Shen, G.: General solution for high dynamic range three-dimensional shape measurement using the fringe projection technique. *Optics and Lasers in Engineering* 59 (2014), pp. 56–71. ISSN: 01438166. DOI: 10.1016/j.optlaseng.2014.03.003.

FENG et al. 2017
Feng, S.; Chen, Q.; Zuo, C.; Asundi, A.: Fast three-dimensional measurements for dynamic scenes with shiny surfaces. *Optics Communications* 382 (2017), pp. 18–27. ISSN: 00304018. DOI: 10.1016/j.optcom.2016.07.057.

FENG et al. 2018
Feng, S.; Zhang, L.; Zuo, C.; Tao, T.; Chen, Q.; Gu, G.: High dynamic range 3D measurements with fringe projection profilometry: a review. *Measurement Science and Technology* 29 (2018) 12, p. 122001. ISSN: 0957-0233. DOI: 10.1088/1361-6501/aae4fb.

FROHM 2008
Frohm, J.: Levels of automation in production systems. PhD thesis. Chalmers University of Technology. 2008. DOI: 10.13140/RG.2.1.2797.7447.

FROHM et al. 2008
Frohm, J.; Lindström Veronica; Winroth, M.; Stahre, J.: Levels of automation in manufacturing. *Ergonomia* (2008).

GAMER et al. 2019
Gamer, T.; Hoernicke, M.; Kloepper, B.; Bauer, R.; Isaksson, A. J.: The Autonomous Industrial Plant -Future of Process Engineering, Operations and Maintenance. *IFAC-PapersOnLine* 52 (2019) 1, pp. 454–460. ISSN: 24058963. DOI: 10.1016/j.ifacol.2019.06.104.

GAO et al. 2015
Gao, W.; Kim, S. W.; Bosse, H.; Haitjema, H.; Chen, Y. L.; Lu, X. D.; Knapp, W.; Weckenmann, A.; Estler, W. T.; Kunzmann, H.: Measurement technologies for precision positioning. *CIRP Annals* 64 (2015) 2, pp. 773–796. DOI: 10.1016/j.cirp.2015.05.009.

GIRSHICK et al. 2014
Girshick, R.; Donahue, J.; Darrell, T.; Malik, J.: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2014, pp. 580–587. ISBN: 978-1-4799-5118-5. DOI: 10.1109/CVPR.2014.81.

GLORIEUX et al. 2020
Glorieux, E.; Franciosa, P.; Ceglarek, D.: Coverage path planning with targetted viewpoint sampling for robotic free-form surface inspection. *Robotics and Computer-Integrated Manufacturing* 61 (2020), p. 101843. DOI: 10.1016/j.rcim.2019.101843.

GONZÁLEZ-BANOS 2001
González-Banos, H.: A randomized art-gallery algorithm for sensor placement. In: *Proceedings of the seventeenth annual symposium on Computational geometry - SCG '01*. Ed. by Souvaine, D. L. New York, New York, USA: ACM Press. 2001, pp. 232–240. ISBN: 158113357X. DOI: 10.1145/378583.378674.

GOODFELLOW et al. 2016
Goodfellow, I.; Bengio, Y.; Courville, A.: *Deep Learning*. MIT Press. 2016. URL: http://www.deeplearningbook.org.

GOSPODNETIĆ et al. 2022
Gospodnetić, P.; Mosbach, D.; Rauhut, M.; Hagen, H.: Viewpoint placement for inspection planning. *Machine Vision and Applications* 33 (2022) 1. ISSN: 1432-1769. DOI: 10.1007/s00138-021-01252-z.

GPY 2012
GPy: GPy: A Gaussian process framework in python. URL: http://github.com/SheffieldML/GPy.

GRANELL et al. 2007
Granell, V.; Frohm, J.; Bruch, J.; Dencker, K.: *Validation of the DYNAMO Methodology for Measuring and Assessing Levels of Automation*. 2007.

GRONLE and OSTEN 2016
Gronle, M.; Osten, W.: View and sensor planning for multi-sensor surface inspection. *Surface Topography: Metrology and Properties* 4 (2016) 2, p. 024009. DOI: 10.1088/2051-672X/4/2/024009.

GÜHRING 2002
Gühring, J.: 3D-Erfassung und Objektrekonstruktion mittels Streifenprojektion. PhD thesis. University of Stuttgart. 2002.

GUPTA et al. 2011
Gupta, M.; Agrawal, A.; Veeraraghavan, A.; Narasimhan, S. G.: Structured light 3D scanning in the presence of global illumination. In: *CVPR 2011*. IEEE. 2011, pp. 713–720. ISBN: 978-1-4577-0394-2. DOI: 10.1109/CVPR.2011.5995321.

HÄGELE et al. 2016
Hägele, M.; Nilsson, K.; Pires, J. N.; Bischoff, R.: Industrial Robotics. In: *Springer Handbook of Robotics*. Ed. by Siciliano, B.; Khatib, O. Cham: Springer International Publishing. 2016, pp. 1385–1422. ISBN: 978-3-319-32550-7. DOI: 10.1007/978-3-319-32552-1_54.

HAMMERSTINGL 2020
Hammerstingl, V.: Steigerung der Rekonfigurationsfähigkeit von Montageanlagen durch Cyber-Physische Feldgeräte. PhD thesis. Technical University of Munich. 2020. DOI: 10.13140/RG.2.2.26067.40489.

HARTL et al. 2021
Hartl, R.; Bachmann, A.; Habedank, J. B.; Semm, T.; Zaeh, M. F.: Process Monitoring in Friction Stir Welding Using Convolutional Neural Networks. *Metals* 11 (2021) 4, p. 535. DOI: 10.3390/met11040535.

HE et al. 2017
He, K.; Gkioxari, G.; Dollar, P.; Girshick, R.: Mask R-CNN. *2017 IEEE International Conference on Computer Vision (ICCV)* (2017). DOI: 10.1109/iccv.2017.322.

HESSE and MALISA 2016
Hesse, S.; Malisa, V.: *Taschenbuch Robotik-Montage-Handhabung*. Carl Hanser Verlag GmbH Co KG. 2016.

HEUSS et al. 2022
Heuss, L.; Gonnermann, C.; Reinhart, G.: An extendable framework for intelligent and easily configurable skills-based industrial robot applications. *The International Journal of Advanced Manufacturing Technology* 120 (2022) 9-10, pp. 6269–6285. ISSN: 0268-3768. DOI: 10.1007/s00170-022-09071-w.

HINTERSTOISSER et al. 2013
Hinterstoisser, S.; Lepetit, V.; Ilic, S.; Holzer, S.; Bradski, G.; Konolige, K.; Navab, N.: Model Based Training, Detection and Pose Estimation of Texture-Less 3D Objects in Heavily Cluttered Scenes. In: *Computer Vision – ACCV 2012*. Vol. 7724. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg. 2013, pp. 548–562. ISBN: 978-3-642-37330-5. DOI: 10.1007/978-3-642-37331-2_42.

HIRZINGER et al. 1994
Hirzinger, G.; Brunner, B.; Dietrich, J.; Heindl, J.: ROTEX-the first remotely controlled robot in space. In: *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. IEEE Comput. Soc. Press. 1994, pp. 2604–2611. ISBN: 0-8186-5330-2. DOI: 10.1109/ROBOT.1994.351121.

HOCKEN 2012
Hocken, R. J.: *Coordinate measuring machines and systems*. 2nd ed. Vol. 76. Manufacturing engineering and materials processing. Boca Raton, London, and New York: CRC Press, Taylor & Francis Group. 2012. ISBN: 9781138076891.

HODAŇ et al. 2016
Hodaň, T.; Matas, J.; Obdržálek, Š.: On Evaluation of 6D Object Pose Estimation. In: *Computer Vision – ECCV 2016 Workshops*. Ed. by Hua, G.; Jégou, H. Vol. 9915. Lecture Notes in Computer Science. Cham: Springer International Publishing. 2016, pp. 606–619. ISBN: 978-3-319-49408-1. DOI: 10.1007/978-3-319-49409-8_52.

HODAN et al. 2017
Hodan, T.; Haluza, P.; Obdrzalek, S.; Matas, J.; Lourakis, M.; Zabulis, X.: T-LESS: An RGB-D Dataset for 6D Pose Estimation of Texture-Less Objects. *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)* (2017). DOI: 10.1109/wacv.2017.103.

HOFFMANN et al. 2021
Hoffmann, M. W.; Drath, R.; Ganz, C.: Proposal for requirements on industrial AI solutions. In: *Machine Learning for Cyber Physical Systems*. Ed. by Beyerer, J.; Maier, A.; Niggemann, O. Vol. 13. Technologien für die intelligente Automation. Berlin, Heidelberg: Springer Berlin Heidelberg. 2021, pp. 63–72. ISBN: 978-3-662-62745-7. DOI: 10.1007/978-3-662-62746-4_7.

HUBEL and WIESEL 1959
Hubel, D. H.; Wiesel, T. N.: Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology* 148 (1959) 3, pp. 574–591. ISSN: 0022-3751. DOI: 10.1113/jphysiol.1959.sp006308.

IMHOFF et al. 2018
Imhoff, F. B.; Schnell, J.; Magaña, A.; Diermeier, T.; Scheiderer, B.; Braun, S.; Imhoff, A. B.; Arciero, R. A.; Beitzel, K.: Single cut distal femoral osteotomy for

correction of femoral torsion and valgus malformity in patellofemoral malalignment - proof of application of new trigonometrical calculations and 3D-printed cutting guides. *BMC musculoskeletal disorders* 19 (2018) 1, p. 215. DOI: 10.1186/s12891-018-2140-5.

ISO 1101 2017
ISO 1101: Geometrical product specifications (GPS). Geometrical tolerancing – Tolerances of form, orientation, location and run-out.

ISO/IEC 25010 2011
ISO/IEC 25010: Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models.

JIANG et al. 2012
Jiang, H.; Zhao, H.; Li, X.: High dynamic range fringe acquisition: A novel 3-D scanning technique for high-reflective surfaces. *Optics and Lasers in Engineering* 50 (2012) 10, pp. 1484–1493. ISSN: 01438166. DOI: 10.1016/j.optlaseng.2011.11.021. URL: https://www.sciencedirect.com/science/article/pii/S0143816612000978.

JING et al. 2017
Jing, W.; Polden, J.; Goh, C. F.; Rajaraman, M.; Lin, W.; Shimada, K.: Sampling-based coverage motion planning for industrial inspection application with redundant robotic system. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 5211–5218. ISBN: 978-1-5386-2682-5. DOI: 10.1109/IROS.2017.8206411.

KABA et al. 2017
Kaba, M. D.; Uzunbas, M. G.; Lim, S. N.: A Reinforcement Learning Approach to the View Planning Problem. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, pp. 5094–5102. DOI: 10.1109/CVPR.2017.541.

KEFERSTEIN and MARXER 2015
Keferstein, C. P.; Marxer, M.: *Fertigungsmesstechnik. Praxisorientierte Grundlagen, moderne Messverfahren / Claus P. Keferstein, Michael Marxer*. Eighth edition. Wiesbaden: Springer Vieweg. 2015. ISBN: 978-3-8348-2583-4.

KEHL et al. 2017
Kehl, W.; Manhardt, F.; Tombari, F.; Ilic, S.; Navab, N.: SSD-6D: Making RGB-based 3D detection and 6D pose estimation great again. *2017 IEEE International Conference on Computer Vision (ICCV)* (2017). DOI: 10.1109/iccv.2017.169.

KENDALL et al. 2015
Kendall, A.; Grimes, M.; Cipolla, R.: PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization. *2015 IEEE International Conference on Computer Vision (ICCV)* (2015). DOI: 10.1109/iccv.2015.336.

KOUTECKÝ et al. 2016
Koutecký, T.; Paloušek, D.; Brandejs, J.: Sensor planning system for fringe projection scanning of sheet metal parts. *Measurement* 94 (2016), pp. 60–70.

KRIEGEL et al. 2015
Kriegel, S.; Rink, C.; Bodenmüller, T.; Suppa, M.: Efficient next-best-scan planning for autonomous 3D surface reconstruction of unknown objects. *Journal of Real-Time Image Processing* 10 (2015) 4, pp. 611–631. ISSN: 1861-8200. DOI: 10.1007/s11554-013-0386-6.

KRITTER et al. 2019
Kritter, J.; Brévilliers, M.; Lepagnot, J.; Idoumghar, L.: On the optimal placement of cameras for surveillance and the underlying set cover problem. *Applied Soft Computing* 74 (2019), pp. 133–153. ISSN: 15684946. DOI: 10.1016/j.asoc.2018.10.025.

KRIZHEVSKY et al. 2017
Krizhevsky, A.; Sutskever, I.; Hinton, G. E.: ImageNet classification with deep convolutional neural networks. *Communications of the ACM* 60 (2017) 6, pp. 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386.

KUBICEK 1976
Kubicek, H.: *Heuristische Bezugsrahmen und heuristisch angelegte Forschungsdesign als Elemente einer Konstruktionsstrategie empirischer Forschung*. Vol. Nr. 16. Arbeitspapier / Institut für Unternehmungsführung im Fachbereich Wirtschaftswissenschaft der Freien Universität Berlin. Berlin: Inst. für Unternehmungsführung im Fachbereich Wirtschaftswiss. d. Freien Univ. 1976. ISBN: 3-88398-016-1.

KUHN and JOHNSON 2013
Kuhn, M.; Johnson, K.: *Applied Predictive Modeling*. New York, NY: Springer New York. 2013. ISBN: 978-1-4614-6848-6. DOI: 10.1007/978-1-4614-6849-3.

LÄMMEL and CLEVE 2012
Lämmel, U.; Cleve, J.: *Künstliche Intelligenz. Mit 51 Tabellen, 43 Beispielen, 118 Aufgaben, 89 Kontrollfragen und Referatsthemen*. 4., aktualisierte Aufl. München: Hanser. 2012. ISBN: 978-3-446-42873-7.

LAURI et al. 2020
Lauri, M.; Pajarinen, J.; Peters, J.; Frintrop, S.: Multi-Sensor Next-Best-View Planning as Matroid-Constrained Submodular Maximization. *IEEE Robotics and Automation Letters* 5 (2020) 4, pp. 5323–5330. DOI: 10.1109/LRA.2020.3007445.

LEE and PARK 2000
Lee, K. H.; Park, H.-p.: Automated inspection planning of free-form shape parts by laser scanning. *Robotics and Computer-Integrated Manufacturing* 16 (2000) 4, pp. 201–210.

LI and KOFMAN 2014
Li, D.; Kofman, J.: Adaptive fringe-pattern projection for image saturation avoidance in 3D surface-shape measurement. *Optics express* 22 (2014) 8, pp. 9887–9901. DOI: 10.1364/OE.22.009887.

LIEBL 2020
Liebl, C.: Systematische Energiedatenerfassung in der Produktion. PhD thesis. Technical University of Munich. 2020.

LIN et al. 2017a
Lin, H.; Gao, J.; Zhang, G.; Chen, X.; He, Y.; Liu, Y.: Review and Comparison of High-Dynamic Range Three-Dimensional Shape Measurement Techniques. *Journal of Sensors* 2017 (2017), pp. 1–11. ISSN: 1687-725X. DOI: 10.1155/2017/9576850.

LIN et al. 2017b
Lin, H.; Gao, J.; Mei, Q.; Zhang, G.; He, Y.; Chen, X.: Three-dimensional shape measurement technique for shiny surfaces by adaptive pixel-wise projection intensity adjustment. *Optics and Lasers in Engineering* 91 (2017), pp. 206–215. ISSN: 01438166. DOI: 10.1016/j.optlaseng.2016.11.015.

LITZENBERGER and KUTZBACH 2016
Litzenberger, G.; Kutzbach, N.: World Robotics 2016 - Industrial Robots.

LIU et al. 2011
Liu, G.-h.; Liu, X.-Y.; Feng, Q.-Y.: 3D shape measurement of objects with high dynamic range of surface reflectivity. *Applied optics* 50 (2011) 23, pp. 4557–4565. DOI: 10.1364/AO.50.004557.

LIU et al. 2016
Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A. C.: SSD: Single Shot MultiBox Detector. In: *Computer Vision – ECCV 2016*. Ed. by Leibe, B.; Matas, J.; Sebe, N.; Welling, M. Lecture Notes in Computer Science. Cham: Springer International Publishing. 2016, pp. 21–37. ISBN: 9783319464480. DOI: 10.1007/978-3-319-46448-0_2.

LÜCK 2004
Lück, W., ed., (2004): *Lexikon der Betriebswirtschaft*. Oldenbourg Wissenschaftsverlag. 2004. ISBN: 9783486815627. DOI: 10.1515/9783486815627.

LUX-GRUENBERG 2020
Lux-Gruenberg, G. F. R.: Selbst-Kalibrierung roboterbasierter Koordinatenmesssysteme. PhD thesis. Technical University of Munich. 2020.

MAGAÑA and REINHART 2017
Magaña, A.; Reinhart, G.: Herstellerneutrale Programmierung von Robotern. *WT Werkstattstechnik* (2017), pp. 594–599. DOI: 10.37544/1436-4980-2017-09.

MAGAÑA et al. 2019
Magaña, A.; Bauer, P.; Reinhart, G.: Concept of a learning knowledge-based system for programming industrial robots. *Procedia CIRP* 79 (2019), pp. 626–631. ISSN: 22128271. DOI: 10.1016/j.procir.2019.02.076.

MAGAÑA et al. 2020a
Magaña, A.; Wu, H.; Bauer, P.; Reinhart, G.: PoseNetwork: Pipeline for the Automated Generation of Synthetic Training Data and CNN for Object Detection, Segmentation, and Orientation Estimation. (2020), pp. 587–594. DOI: 10.1109/ETFA46521.2020.9212064.

MAGAÑA et al. 2020b
Magaña, A.; Gebel, S.; Bauer, P.; Reinhart, G.: Knowledge-Based Service-Oriented System for the Automated Programming of Robot-Based Inspection Systems. (2020), pp. 1511–1518. DOI: 10.1109/ETFA46521.2020.9212033.

MAGAÑA et al. 2023a
Magaña, A.; Dirr, J.; Bauer, P.; Reinhart, G.: Viewpoint Generation Using Feature-Based Constrained Spaces for Robot Vision Systems. *Robotics* 12 (2023) 4, p. 108. DOI: 10.3390/robotics12040108.

MAGAÑA et al. 2023b
Magaña, A.; Vlaeyen, M.; Haitjema, H.; Bauer, P.; Schmucker, B.; Reinhart, G.: Viewpoint Planning for Range Sensors Using Feature Cluster Constrained Spaces for Robot Vision Systems. *Sensors* 23 (2023) 18, p. 7964. DOI: 10.3390/s23187964.

MAGAÑA et al. 2023c
Magaña, A.; Schneider, L.; Benker, M.; Altmann, T.; Bauer, P.; Reinhart, G.: Exposure Time and Point Cloud Quality Prediction for Active 3D Imaging Sensors using Gaussian Processes. *Production Engineering* (2023). ISSN: 0944-6524. DOI: 10.1007/s11740-023-01240-4.

MAGAÑA et al. 2023d
Magaña, A.; Dirr, J.; Bauer, P.; Reinhart, G.: Viewpoint Generation using Feature-Based Constrained Spaces for Robot Vision Systems. (2023). DOI: 10.48550/arXiv.2306.06969.

MAVRINAC and CHEN 2013
Mavrinac, A.; Chen, X.: Modeling Coverage in Camera Networks: A Survey. *International Journal of Computer Vision* 101 (2013) 1, pp. 205–226. ISSN: 0920-5691. DOI: 10.1007/s11263-012-0587-7.

MAVRINAC et al. 2015
Mavrinac, A.; Chen, X.; Alarcon-Herrera, J. L.: Semiautomatic Model-Based View Planning for Active Triangulation 3-D Inspection Systems. *IEEE/ASME Transactions on Mechatronics* 20 (2015) 2, pp. 799–811. ISSN: 1083-4435. DOI: 10.1109/TMECH.2014.2318729.

MITCHELL 1997
Mitchell, T. M.: *Machine Learning*. McGraw-Hill series in computer science. New York: McGraw-Hill. 1997. ISBN: 0070428077.

MITTAL and DAVIS 2007
Mittal, A.; Davis, L. S.: A General Method for Sensor Planning in Multi-Sensor Systems: Extension to Random Occlusion. *International Journal of Computer Vision* 76 (2007) 1, pp. 31–52. ISSN: 0920-5691. DOI: 10.1007/s11263-007-0057-9.

MOHAMMADIKAJI 2019
Mohammadikaji, M.: Simulation-based Planning of Machine Vision Inspection Systems with an Application to Laser Triangulation. (2019).

MOSBACH et al. 2021
Mosbach, D.; Gospodnetić, P.; Rauhut, M.; Hamann, B.; Hagen, H.: Feature-Driven Viewpoint Placement for Model-Based Surface Inspection. *Machine Vision and Applications* 32 (2021) 1. ISSN: 1432-1769. DOI: 10.1007/s00138-020-01116-y.

MÜLLER and KUTZBACH 2019
Müller, C.; Kutzbach, N.: World Robotics 2019 - Industrial Robots. Germany.

MÜLLER and ARENS 2010
Müller, J.; Arens, M.: Human pose estimation with implicit shape models. In: *Proceedings of the first ACM international workshop on Analysis and retrieval of tracked events and motion in imagery streams - ARTEMIS '10*. Ed. by Doulamis, A.; Gonzàlez, J.; Bertini, M.; van Gool, L.; Matsatsinis, N.; Moeslund, T. B.; Nixon, M.; Wang, L. New York, New York, USA: ACM Press. 2010, p. 9. ISBN: 9781450301633. DOI: 10.1145/1877868.1877873.

MURPHY 2012
Murphy, K. P.: *Machine learning. A probabilistic perspective*. Adaptive computation and machine learning series. Cambridge MA: MIT Press. 2012. ISBN: 9780262018029.

MURPHY 2022
— *Probabilistic machine learning. An introduction*. Adaptive computation and machine learning series. Cambridge Massachusetts: The MIT Press. 2022. ISBN: 0262046822.

NAYAR et al. 1991
Nayar, S. K.; Ikeuchi, K.; Kanade, T.: Surface reflection: physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and machine intelligence* 13 (1991) 7, pp. 611–634. ISSN: 01628828. DOI: 10.1109/34.85654.

PALOUSEK et al. 2015
Palousek, D.; Omasta, M.; Koutny, D.; Bednar, J.; Koutecky, T.; Dokoupil, F.:

Effect of matte coating on 3D optical measurement accuracy. *Optical Materials* 40 (2015), pp. 1–9. ISSN: 0925-3467. DOI: 10.1016/j.optmat.2014.11.020.

PAN et al. 2012
Pan, Z.; Polden, J.; Larkin, N.; van Duin, S.; Norrish, J.: Recent progress on programming methods for industrial robots. *Robotics and Computer-Integrated Manufacturing* 28k (2012) 2, pp. 87–94.

PARK et al. 2006
Park, J.; Bhat, P. C.; Kak, A. C.: A Look-up Table Based Approach for Solving the Camera Selection Problem in Large Camera Networks. (2006).

PEUZIN-JUBERT et al. 2021
Peuzin-Jubert, M.; Polette, A.; Nozais, D.; Mari, J.-L.; Pernot, J.-P.: Survey on the View Planning Problem for Reverse Engineering and Automated Control Applications. *Computer-Aided Design* 141 (2021), p. 103094. DOI: 10.1016/j.cad.2021.103094.

PFEIFER 2002
Pfeifer, T.: *Production Metrology*. Munchen and Berlin: Oldenbourg Wissenschaftsverlag and Walter de Gruyter GmbH & Co. KG. 2002. ISBN: 3486258850.

PFEIFER and SCHMITT 2010
Pfeifer, T.; Schmitt, R.: *Fertigungsmesstechnik*. 3rd ed. München: Oldenbourg. 2010. ISBN: 3-48659-202-5.

PINHEIRO et al. 2015
Pinheiro, P. O.; Collobert, R.; Dollar, P.: Learning to segment object candidates. In: *Advances in Neural Information Processing Systems*. 2015, pp. 1990–1998.

PITO 1999
Pito, R.: A solution to the next best view problem for automated surface acquisition. *IEEE Transactions on Pattern Analysis and machine intelligence* 21 (1999) 10, pp. 1016–1030. ISSN: 01628828. DOI: 10.1109/34.799908.

QUIGLEY et al. 2009
Quigley, M.; Conley, K.; Gerkey, B.; Faust, J.; Foote, T.; Leibs, J.; Wheeler, R.; Ng, A. Y., et al.: ROS: an open-source Robot Operating System. In: *ICRA workshop on open source software*. Vol. 3. 2009, p. 5.

RAFFAELI et al. 2013
Raffaeli, R.; Mengoni, M.; Germani, M.; Mandorli, F.: Off-line view planning for the inspection of mechanical parts. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 7 (2013) 1, pp. 1–12. ISSN: 1955-2513. DOI: 10.1007/s12008-012-0160-1.

RASMUSSEN and WILLIAMS 2006
Rasmussen, C. E.; Williams, C. K. I.: *Gaussian processes for machine learning*.

Adaptive computation and machine learning. Cambridge, Mass. and London: MIT. 2006. ISBN: 978-0-262-18253-9.

REDMON et al. 2016
Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016). DOI: 10.1109/cvpr.2016.91.

REED and ALLEN 2000
Reed, M. K.; Allen, P. K.: Constraint-based sensor planning for scene modeling. *IEEE Transactions on Pattern Analysis and machine intelligence* 22 (2000) 12, pp. 1460–1467. ISSN: 01628828. DOI: 10.1109/34.895979.

REED 1998
Reed, M.: Solid model acquisition from range imagery. PhD thesis. Universität Stuttgart. 1998.

REINHART et al. 2018
Reinhart, G.; Magaña Flores, A.; Zwicker, C.: *Industrieroboter. Planung - Integration - Trends: ein Leitfaden für KMU*. 1. Auflage. Ein Fachbuch von Konstruktionspraxis, Elektrotechnik, MM MaschinenMarkt. Würzburg: Vogel Communications Group. 2018. ISBN: 3834334014.

REN et al. 2017
Ren, S.; He, K.; Girshick, R.; Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and machine intelligence* (2017) 6. DOI: 10.1109/TPAMI.2016.2577031.

SALAHIEH et al. 2014
Salahieh, B.; Chen, Z.; Rodriguez, J. J.; Liang, R.: Multi-polarization fringe projection imaging for high dynamic range objects. *Optics express* 22 (2014) 8, pp. 10064–10071. DOI: 10.1364/OE.22.010064.

SANSONI et al. 2009
Sansoni, G.; Trebeschi, M.; Docchio, F.: State-of-The-Art and Applications of 3D Imaging Sensors in Industry, Cultural Heritage, Medicine, and Criminal Investigation. *Sensors (Basel, Switzerland)* 9 (2009) 1, pp. 568–601. DOI: 10.3390/s90100568.

SCHMITT et al. 2011
Schmitt, R.; Nisch, S.; Heizmann, M.; Bosse, H.; Imkamp, D.: Production Metrology–Future Trends and Challenges. In: *Proceedings of the 10th International Symposium on Measurement Technology and Intelligent Instruments (ISMTII-2011) Daejeon, S. Korea 29th June–2nd July*. 2011.

SCHMITT et al. 2016
Schmitt, R. H.; Peterek, M.; Morse, E.; Knapp, W.; Galetto, M.; Härtig, F.; Goch, G.; Hughes, B.; Forbes, A.; Estler, W. T.: Advances in Large-Scale Metrology

– Review and future trends. *CIRP Annals* 65 (2016) 2, pp. 643–665. DOI: 10.1016/j.cirp.2016.05.002.

SCOTT 2002
Scott, W. R.: Performance-oriented view planning for automated object reconstruction. PhD thesis. University of Ottawa. 2002.

SCOTT 2009
— Model-based view planning. *Machine Vision and Applications* 20 (2009) 1, pp. 47–69. ISSN: 1432-1769. DOI: 10.1007/s00138-007-0110-2.

SCOTT et al. 2003
Scott, W. R.; Roth, G.; Rivest, J.-F.: View planning for automated three-dimensional object reconstruction and inspection. *ACM Computing Surveys (CSUR)* 35 (2003) 1, pp. 64–96.

SHETTY et al. 2021
Shetty, A. K.; Saha, I.; Sanghvi, R. M.; Save, S. A.; Patel, Y. J.: A Review: Object Detection Models. In: *2021 6th International Conference for Convergence in Technology (I2CT)*. IEEE. 2021, pp. 1–8. ISBN: 978-1-7281-8876-8. DOI: 10.1109/I2CT51068.2021.9417895.

SHI et al. 2006
Shi, Q.; Xi, N.; Sheng, W.; Chen, Y.: Development of dynamic inspection methods for dimensional measurement of automotive body parts. In: *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006*. IEEE. 2006, pp. 315–320. ISBN: 0-7803-9505-0. DOI: 10.1109/ROBOT.2006.1641730.

SKOCAJ and LEONARDIS 2000
Skocaj, D.; Leonardis, A.: Range image acquisition of objects with non-uniform albedo using structured light range sensor. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. IEEE Comput. Soc. 2000, pp. 778–781. ISBN: 0-7695-0750-6. DOI: 10.1109/ICPR.2000.905506.

SONG et al. 2017
Song, Z.; Jiang, H.; Lin, H.; Tang, S.: A high dynamic range structured light means for the 3D measurement of specular surface. *Optics and Lasers in Engineering* 95 (2017), pp. 8–16. ISSN: 01438166. DOI: 10.1016/j.optlaseng.2017.03.008.

STENMARK and MALEC 2015
Stenmark, M.; Malec, J.: Knowledge-based instruction of manipulation tasks for industrial robotics. *Robotics and Computer-Integrated Manufacturing* 33 (2015), pp. 56–67. DOI: 10.1016/j.rcim.2014.07.004.

STÖSSEL et al. 2004
Stößel, D.; Hanheide, M.; Sagerer, G.; Krüger, L.; Ellenrieder, M.: Feature and viewpoint selection for industrial car assembly. In: *Joint Pattern Recognition Symposium*. 2004, pp. 528–535.

Su et al. 2021
Su, Y.-H.; Huang, K.; Hannaford, B.: Multicamera 3D Viewpoint Adjustment for Robotic Surgery via Deep Reinforcement Learning. *Journal of Medical Robotics Research* 06 (2021) 01n02, p. 2140003. ISSN: 2424-905X. DOI: 10.1142/S2424905X21400031.

Sun et al. 2017
Sun, X.; Liu, Y.; Yu, X.; Wu, H.; Zhang, N.: Three-Dimensional Measurement for Specular Reflection Surface Based on Reflection Component Separation and Priority Region Filling Theory. *Sensors (Basel, Switzerland)* 17 (2017) 1. DOI: 10.3390/s17010215.

Sundermeyer et al. 2018
Sundermeyer, M.; Marton, Z.-C.; Durner, M.; Brucker, M.; Triebel, R.: Implicit 3d orientation learning for 6d object detection from rgb images. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018, pp. 699–715.

Szegedy et al. 2015
Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A.: Going deeper with convolutions. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. 2015, pp. 1–9. ISBN: 978-1-4673-6964-0. DOI: 10.1109/CVPR.2015.7298594.

Szeliski 2022
Szeliski, R.: *Computer Vision: Algorithms and Applications*. 2nd ed. Springer. 2022. ISBN: 978-3030343712.

Tarabanis et al. 1995a
Tarabanis, K. A.; Tsai, R. Y.; Allen, P. K.: The MVP sensor planning system for robotic vision tasks. *IEEE transactions on Robotics and Automation* 11 (1995) 1, pp. 72–85. ISSN: 1042-296X. DOI: 10.1109/70.345939.

Tarabanis and Tsai 1991
Tarabanis, K.; Tsai, R. Y.: Computing viewpoints that satisfy optical constraints. In: *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1991, pp. 152–158. DOI: 10.1109/CVPR.1991.139678.

Tarabanis et al. 1995b
Tarabanis, K. A.; Allen, P. K.; Tsai, R. Y.: A survey of sensor planning in computer vision. *IEEE transactions on Robotics and Automation* 11 (1995) 1, pp. 86–104. ISSN: 1042-296X. DOI: 10.1109/70.345940.

Tarbox and Gottschlich 1994
Tarbox, G. H.; Gottschlich, S. N.: IVIS: An integrated volumetric inspection system. *Computer Vision and Image Understanding* 61 (1994) 3, pp. 430–444. DOI: 10.1109/CADVIS.1994.284498.

TARBOX and GOTTSCHLICH 1995
— Planning for complete sensor coverage in inspection. *Computer Vision and Image Understanding* 61 (1995) 1, pp. 84–111. DOI: 10.1006/cviu.1995.1007.

TEKOUO MOUTCHIHO 2012
Tekouo Moutchiho, W. B.: A New Programming Approach for Robot-based Flexible Inspection Systems. PhD thesis. Technical University of Munich. 2012.

TENSORFLOW DEVELOPERS 2023
TensorFlow Developers: TensorFlow. DOI: 10.5281/zenodo.4724125. URL: https://github.com/tensorflow/tensorflow (visited on 12/03/2023).

THE BOSTON CONSULTING GROUP 2015
The Boston Consulting Group: Takeoff in Robotics Will Power the Next Productivity Surge in Manufacturing - Selected Highlights.

TIGHINEANU et al. 2021
Tighineanu, P.; Skubch, K.; Baireuther, P.; Reiss, A.; Berkenkamp, F.; Vinogradska, J.: Transfer Learning with Gaussian Processes for Bayesian Optimization. DOI: 10.48550/arXiv.2111.11223.

TREMBLAY et al. 2018a
Tremblay, J.; To, T.; Sundaralingam, B.; Xiang, Y.; Fox, D.; Birchfield, S.: Deep object pose estimation for semantic robotic grasping of household objects. (2018). DOI: 10.48550/arXiv.1809.10790.

TREMBLAY et al. 2018b
Tremblay, J.; Prakash, A.; Acuna, D.; Brophy, M.; Jampani, V.; Anil, C.; To, T.; Cameracci, E.; Boochoon, S.; Birchfield, S.: Training deep networks with synthetic data: Bridging the reality gap by domain randomization. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018, pp. 969–977.

TRUCCO et al. 1997
Trucco, E.; Umasuthan, M.; Wallace, A. M.; Roberto, V.: Model-based planning of optimal sensor placements for inspection. *IEEE transactions on Robotics and Automation* 13 (1997) 2, pp. 182–194. ISSN: 1042-296X. DOI: 10.1109/70.563641.

TSIOUVARAS et al. 2019
Patent: DE102019105064A1 pat. (03.09.2020). 2019. Tsiouvaras, N.; Daub, R.; Bachmann, F.; Semm, T.; Richter, L.; Magaña, A.; Metkar, A.: „Prüfvorrichtung und Verfahren zum Testen einer Speicherzelle“.

ULRICH et al. 2015
Ulrich, M.; Forstner, A.; Reinhart, G.: High-accuracy 3D image stitching for robot-based inspection systems. In: *2015 IEEE International Conference on Image Processing (ICIP)*. 2015, pp. 1011–1015. DOI: 10.1109/ICIP.2015.7350952.

ULRICH 2018
Ulrich, M.: 3D-Image-Stitching für roboterbasierte Messsysteme. PhD thesis. Technical University of Munich. 2018.

ULRICH and HILL 1976
Ulrich, P.; Hill, W.: Wissenschaftstheoretische Grundlagen der Betriebswirtschaftslehre. (1976).

UNITED NATIONS 2022
United Nations: Transforming our world: the 2030 Agenda for Sustainable Development. In: *Sustainable Development Goals*. Ed. by Huck, W. Nomos Verlagsgesellschaft mbH & Co. KG. 2022. ISBN: 9783748902065. DOI: 10.5771/9783748902065-653.

UNITY TECHNOLOGIES 2022
Unity Technologies: Unity. URL: https://unity.com/ (visited on 12/03/2023).

VAGIA et al. 2016
Vagia, M.; Transeth, A. A.; Fjerdingen, S. A.: A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? *Applied ergonomics* 53 Pt A (2016), pp. 190–202. DOI: 10.1016/j.apergo.2015.09.013.

VIOLA and JONES 2001
Viola, P.; Jones, M.: Robust real-time face detection. In: *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*. IEEE Comput. Soc. 2001, p. 747. ISBN: 0-7695-1143-0. DOI: 10.1109/ICCV.2001.937709.

VOEGELE and SOMMER 2012
Voegele, A. A.; Sommer, L.: *Kosten- und Wirtschaftlichkeitsrechnung für Ingenieure. Kostenmanagement im Engineering*. München: Hanser, Carl. 2012. ISBN: 978-3-446-42975-8. DOI: Arno.

VOGL 2009
Vogl, W.: *Eine interaktive räumliche Benutzerschnittstelle für die Programmierung von Industrierobotern*. Vol. 228. Herbert Utz Verlag. 2009.

WADDINGTON and KOFMAN 2010
Waddington, C.; Kofman, J.: Sinusoidal fringe-pattern projection for 3-D surface measurement with variable illuminance. In: *2010 International Symposium on Optomechatronic Technologies*. IEEE. 2010, pp. 1–5. ISBN: 978-1-4244-7684-8. DOI: 10.1109/ISOT.2010.5687389.

WADDINGTON and KOFMAN 2014a
— Camera-independent saturation avoidance in measuring high-reflectivity-variation surfaces using pixel-wise composed images from projected patterns of different maximum gray level. *Optics Communications* 333 (2014), pp. 32–37. ISSN: 00304018. DOI: 10.1016/j.optcom.2014.07.039.

WADDINGTON and KOFMAN 2014b
— Modified sinusoidal fringe-pattern projection for variable illuminance in phase-shifting three-dimensional surface-shape metrology. *Optical Engineering* 53 (2014) 8, p. 084109. ISSN: 0091-3286.

WALDRON and SCHMIEDELER 2016
Waldron, K. J.; Schmiedeler, J.: Kinematics. In: *Springer Handbook of Robotics*. Ed. by Siciliano, B.; Khatib, O. Cham: Springer International Publishing. 2016, pp. 11–36. ISBN: 978-3-319-32550-7. DOI: 10.1007/978-3-319-32552-1_2.

WATSON and SCHEIDT 2005
Watson, D. P.; Scheidt, D. H.: *Autonomous Systems*. Johns Hopkins Applied Physics Laboratory Technical Digest. 2005.

WEBER and SEEBERG 2020
Weber, R.; Seeberg, P.: *KI in der Industrie. Grundlagen, anwendungen, perspektiven;grundlagen, anwendungen, perspektiven*. Cincinnati: Hanser. 2020. ISBN: 978-3-446-46427-8.

WECKENMANN 2012
Weckenmann, A.: *Koordinatenmesstechnik. Flexible Strategien für funktions- und fertigungsgerechtes Prüfen*. Carl Hanser Verlag GmbH & Co. KG. 2012. ISBN: 978-3-446-40739-8.

XIANG et al. 2018
Xiang, Y.; Schmidt, T.; Narayanan, V.; Fox, D.: PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes. *Robotics: Science and Systems XIV* (2018). DOI: 10.15607/RSS.2018.XIV.019.

ZHANG et al. 2021
Zhang, H.; Eastwood, J.; Isa, M.; Sims-Waterhouse, D.; Leach, R.; Piano, S.: Optimisation of camera positions for optical coordinate measurement based on visible point analysis. *Precision Engineering* 67 (2021), pp. 178–188. ISSN: 01416359. DOI: 10.1016/j.precisioneng.2020.09.016.

ZHANG 2020
Zhang, S.: Rapid and automatic optimal exposure control for digital fringe projection technique. *Optics and Lasers in Engineering* 128 (2020), p. 106029. ISSN: 01438166. DOI: 10.1016/j.optlaseng.2020.106029.

ZHANG and YAU 2009
Zhang, S.; Yau, S.-T.: High dynamic range scanning technique. *Optical Engineering* 48 (2009) 3, p. 033604. ISSN: 0091-3286. DOI: 10.1117/1.3099720.

ZHANG et al. 2017
Zhang, Y.; Song, S.; Yumer, E.; Savva, M.; Lee, J.-Y.; Jin, H.; Funkhouser, T.: Physically-Based Rendering for Indoor Scene Understanding Using Convolutional Neural Networks. In: *2017 IEEE Conference on Computer Vision and Pattern*

Recognition (CVPR). IEEE. 2017, pp. 5057–5065. ISBN: 978-1-5386-0457-1. DOI: 10.1109/CVPR.2017.537.

ZHAO et al. 2014
Zhao, H.; Liang, X.; Diao, X.; Jiang, H.: Rapid in-situ 3D measurement of shiny object based on fast and high dynamic range digital fringe projector. *Optics and Lasers in Engineering* 54 (2014), pp. 170–174. ISSN: 01438166. DOI: 10.1016/j.optlaseng.2013.08.002.

ZHENG et al. 2019
Zheng, Y.; Wang, Y.; Suresh, V.; Li, B.: Real-time high-dynamic-range fringe acquisition for 3D shape measurement with a RGB camera. *Measurement Science and Technology* 30 (2019) 7, p. 075202. ISSN: 0957-0233. DOI: 10.1088/1361-6501/ab0ced.

ZHOU et al. 2018
Zhou, Q.-Y.; Park, J.; Koltun, V.: Open3D: A Modern Library for 3D Data Processing. (2018). DOI: 10.48550/arXiv.1801.09847.

ZHOU et al. 2016
Zhou, Q.; Grinspun, E.; Zorin, D.; Jacobson, A.: Mesh arrangements for solid geometry. *ACM Transactions on Graphics* 35 (2016) 4, pp. 1–15. ISSN: 0730-0301. DOI: 10.1145/2897824.2925901.

ZHOU and WU 2011
Zhou, Y.; Wu, Y.: Analyses on Influence of Training Data Set to Neural Network Supervised Learning Performance. In: *Advances in Computer Science, Intelligent System and Environment*. Ed. by Kacprzyk, J.; Jin, D.; Lin, S. Vol. 106. Advances in Intelligent and Soft Computing. Berlin, Heidelberg: Springer Berlin Heidelberg. 2011, pp. 19–25. ISBN: 978-3-642-23752-2. DOI: 10.1007/978-3-642-23753-9_4.

ZHU et al. 2022
Zhu, Y.; Li, M.; Yao, W.; Chen, C.: A Review of 6D Object Pose Estimation. In: *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*. IEEE. 2022, pp. 1647–1655. ISBN: 978-1-6654-2207-9. DOI: 10.1109/ITAIC54216.2022.9836663.

List of Figures

List of Tables

Appendix A

Appendix

A.1 Fundamentals

A.1.1 Robot Kinematic Model

The workspace of a six-axis industrial robot is considered to be a subset of the special Euclidean[1] $\mathcal{W}_r \subseteq SE(3)$. This topological space comprises all reachable poses[2] by the robot, thus $\boldsymbol{p}_r \in \mathcal{W}_r$. The nominal poses and joint configurations of a robot can be calculated, if its kinematic model is known. The robot's kinematic model is described by its DH parameters. The DH parameters refer to a set of four parameters (link length and twist and joint offset and angle) that describe the geometric relationships between two consecutive joints of a serial kinematic chain. The DH parameters are used to describe a homogeneous transformation matrix that combines translational and rotational transformation matrices. (HESSE and MALISA 2016, pp. 204–208; WALDRON and SCHMIEDELER 2016, pp. 26–27)

The robot pose for a set of DH parameters is given by successively multiplying the homogeneous transformation matrices of all joints. The robot pose is usually given as the transformation between the robot's last (here sixth) joint C_{j_6} and the robot base C_r as follows:

$$\boldsymbol{p}_r^{j_6} := \boldsymbol{T}_r^{j_6} = \boldsymbol{T}_r^{j_1}(DH_1) \cdot \ldots \cdot \boldsymbol{T}_{j_5}^{j_6}(DH_6). \tag{A.1}$$

Additionally, the robot pose can be given as the homogeneous transformation between the robot base and another reference coordinate system such as the tool center point (TCP) of the end-effector C_{tcp}. Assuming that the rigid transformation between the last joint and the reference coordinate system of the TCP,

[1] The special Euclidean $SE(3) = \mathbb{R}^3 \times SO(3)$ models the product space resulting from the Euclidean space in $\mathbb{R}^3$ and the rotation space denoted as the special orthogonal $SO(3) \subset \mathbb{R}^{3x3}$ (WALDRON and SCHMIEDELER 2016, p. 20).

[2] The pose of a rigid body $\boldsymbol{p} \in SE(3)$, which is an element of the special Euclidean, is fully defined by a translational component $\boldsymbol{t} \in \mathbb{R}^3$ and a rotational component given by a rotation matrix $\boldsymbol{R} \in \mathbb{R}^{3x3}$.

denoted as $\mathbf{T}_{j_6}^{tcp}$, is known, the TCP pose is given as follows:

$$\boldsymbol{p}_r^{tcp} := \boldsymbol{T}_r^{tcp} = \boldsymbol{T}_r^{j_6} \cdot \boldsymbol{T}_{j_6}^{tcp}. \tag{A.2}$$

Since the joint angles are a subset of the DH parameters, the resulting TCP pose of a nominal joint configuration is seamlessly computed using Eq. A.2. This mathematical calculation, known as forward kinematic transformation, is performed very efficiently as it mainly relies on the multiplication of matrices. In contrast, determining the joint angles for a given sensor position is known as the inverse kinematic problem and requires solving nonlinear equations. A more comprehensive derivation of the robot kinematic, forward, and inverse kinematic models is given by WALDRON and SCHMIEDELER (2016, pp. 28–31).

A.1.2 Range Sensor

A.1.2.1 Active Projection

Active range imaging sensors such as *structured light sensor*s (SLSs) consist of at least one camera and a digital projector (see the right graphic of Fig. 1.2). The acquisition principle of depth information is based on active triangulation[3], which assumes that a surface point being represented by a camera image pixel can be matched to a projected point of the light source. In this case, the active source unit projects an encoded sequence of at least three different fringe patterns (e.g., binary values, gray or color values) onto the object surface. For each projection, the camera captures one image at a time. In a next step, the correspondences between projected patterns and image pixels are found using different decoding techniques depending on the projected pattern. (BEYERER et al. 2016, pp. 263–268) Figure A.1 illustrates some of these steps showing a simplified decoding process for a 1D binary projected pattern.

A.1.2.2 Measurement Volume

The measurement volume describes the workspace in which the sensor can measure at standstill in compliance with the specifications provided by the manufacturer (DIN EN ISO 10360-13 2023) . This workspace is described by a set of different sensor imaging parameters, such as the depth of field d_s and the horizontal and vertical field of view (FOV) angles θ_s^x and ψ_s^y (COWAN and KOVESI 1988; TARABANIS et al. 1995a). In addition, some sensor manufacturers provide the dimensions and locations of the near h_s^{near}, middle h_s^{middle}, and far h_s^{far} viewing planes. The measurement volume can be straightforwardly

[3]Triangulation refers to a method of measuring distances using trigonometric principles to measure planar triangles BEYERER et al. (2016, p. 255). A triangle can be defined by two angles and the length of one side, known as the baseline. Using this information, other geometric relationships, such as the distance between a surface point and the camera, can be estimated through trigonometric calculations.

calculated based on the kinematic relationships of the sensor and the imaging parameters. Figure A.2 visualizes the geometrical relationships and frustum shape of the measurement volume.

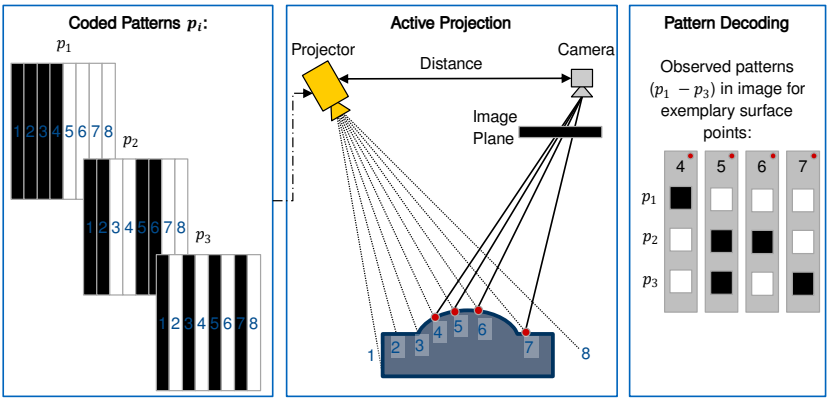


Figure A.1: Simplified acquisition principle in 2D of structured light sensors for range imaging to decode pixel information from a coded projected pattern (BEYERER et al. 2016, pp. 264–265)

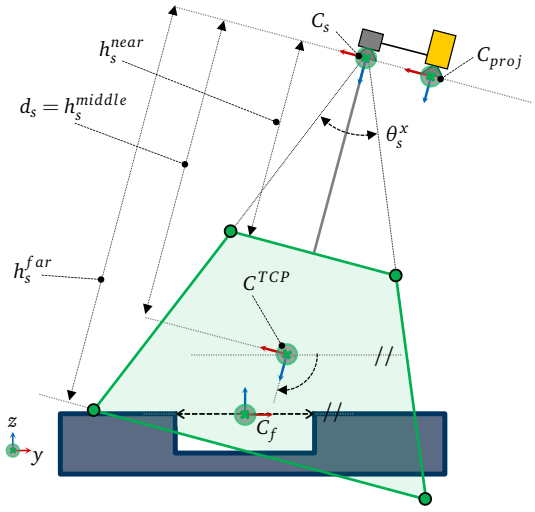


Figure A.2: Detailed kinematic and imaging model of the sensor in the x-z plane. The measurement volume is spanned by the imaging parameters of the sensor ($d_s, h_s^{near}, h_s^{far}, \theta_s^x$).

A.1.3 Machine Learning

Machine learning (ML) is understood as a set of methods that can automatically learn patterns from empirical data and then use the inferred correlations to predict unknown events or make decisions under uncertainty (MURPHY 2012, p.1). MITCHELL (1997, p. 2) introduced a widely quoted definition to describe this type of algorithms:

> *"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."*

This subsection presents the basics of machine learning to provide a general understanding of the mathematical foundations of some of the solution modules proposed in this work. Moreover, the current thesis proposes the use of two different supervised learning methods, *artificial neural network* (ANN) and *Gaussian processes* (GPs), to improve the automation of different subtasks for programming *robot-based optical 3D measuring systems* (ROMSs). The mathematical foundations of these models are more comprehensibly described in the following subsections.

A.1.3.1 Supervised Learning

According to the learning mechanism, machine learning algorithms are classified into the following three main categories depending on what kind of experience E they are allowed to have during the learning stage: supervised learning, unsupervised learning, and reinforcement learning (MURPHY 2012, p. 2; GOODFELLOW et al. 2016, p. 104).

Supervised learning refers to the process of learning input-output mappings from an empirical dataset (RASMUSSEN and WILLIAMS 2006, p.1). Since the input and output data must be labeled in advance, supervised models tend to perform better than unsupervised and reinforcement learning models. For this reason, supervised learning remains one of the most widely used forms of machine learning in real applications (MURPHY 2022, p.1; CHOLLET 2018, p. 94).

Supervised learning methods are further categorized between classification and regression problems depending on the type of task T and the output variable's characteristics. For example, in the context of production tasks, classification algorithms can be used to identify different objects in an image for bin-picking tasks or estimate the wear condition of a workpiece or a machine for predictive maintenance. On the contrary, regression algorithms are employed to predict continuous output values. This model type is particularly helpful, e.g., for the optimization of process parameters, time to failures analyses, or prediction of product demand. Further industrial applications examples are provided by WEBER and SEEBERG (2020).

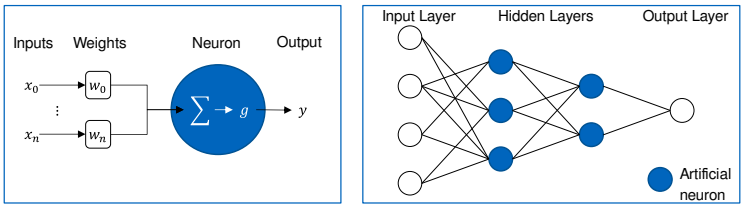
A.1.3.2 Artificial Neural Networks

The elementary component of ANNs is given by an artificial neuron, which represents a mathematical function based on a biological neuron model. A neuron is given a set of n inputs $\{x_0, \ldots, x_n\}$, which are weighted using a set of factors denoted as $\{w_0, \ldots, w_n\}$. In the next step, all weighted inputs are summed and given to an activation function g (e.g., sigmoid, step, sign, linear) to compute an output y. The output of an artificial neuron is calculated as follows:

$$y = g\left(\sum_{i=0}^{n} x_i w_i\right) \tag{A.3}$$

In general, the function g can be arbitrarily chosen to express any complex correlations between the inputs and outputs of a model. Fig. A.3a depicts a simplified representation of a neuron and its components.

The performance of neuron-based approaches is not attributed to the complexity of their functions, but to their architecture. An ANN can be built from one or multiple layers of interconnected neurons with simple functions to solve complex problems (LÄMMEL and CLEVE 2012, p. 192). The inputs, outputs, and topology structure of the neuron layers define the characteristics and architecture of an ANN. The simplest form of ANNs are feed-forward neural networks. Information flows in one direction and each neuron from a neuron layer is connected to all neurons in the successive layer (BASLER 2021, p. 15). Due to this characteristic, feed-forward networks are commonly denoted as fully connected networks. The structure of such a network comprising an input layer, two neuron layers, and an output layer is illustrated in Fig. A.3b.



(a) Artificial neuron model (b) Feed-forward Neural Network

Figure A.3: Simplified representation of an artificial neuron (a) and the architecture of a feed-forward neural network (b) composed by multiple artificial neurons

Before using an ANN for inference, the model parameters (e.g., weights, function hyperparameters) must be selected appropriately in advance. The iterative adjustment of these model hyperparameters using a *training dataset* is referred to as the *learning process* (CHOLLET 2018, p. 10). This process is categorized as an optimization problem, which seeks to minimize the prediction errors using a loss function to increase the model's performance P (BAUER and WARSCHAT 2021, p. 59). To this end, the model is trained based on an experience E that

is represented by a dataset, $\mathcal{D} = \{(x_j, y_j)|j = 1, \ldots, m\}$, with m observations of labeled input-output pairs (MURPHY 2022, p. 2).

Deep neural networks

Deep neural networks (DNNs) are a subcategory of ANNs that consist of a complex model structure with multiple hidden layers. The hidden layers may have a different number of neurons and are successively interconnected with each other. This kind of structure facilitates finer and more flexible filtering of relevant information of the raw data at different abstraction levels for solving more complex problems. CHOLLET (2018, p. 10) defines deep learning as a multistage process to learn data representations.

Due to the large number of weights between layers and neurons that compose a DNN, these models require a high number of observations to be trained. For this reason, training DNNs is often considered a computationally expensive process that demands a large amount of training data for optimizing all model hyperparameters. However, due to the advancements in computational power and decreasing costs in the recent years, DNNs have demonstrated to be valuable and efficient in tackling a wide range of challenging problems (e.g., image classification, character recognition, natural language processing) within diverse applications fields such as neuroscience, pharmaceutical, manufacturing, logistics, and transportation, among others (CHOLLET 2018, p. 94; GOODFELLOW et al. 2016, pp. 23–26).

Convolutional neural networks

Convolutional neural networks (CNNs) represent a special variant of DNNs. The origins of CNNs architectures are traced back to the studies of HUBEL and WIESEL (1959), which investigated the biological vision system of cats (GOODFELLOW et al. 2016, pp. 365–372). Using a similar functionality and structure to a biological vision system, CNNs can extract and interpret patterns from 2D images at low and high abstraction levels. Three types of layers are fundamentally used to accomplish this: *convolutional*, *pooling*, and *fully-connected*.

The function and structure of CNNs is better understood by explaining the inference procedure based on the example of an image classification task to detect an object within an image. A simplified architecture of a CNN for object detection application is depicted in Fig. A.4. The main operations are explained as follows.

In the first step, the input image is represented by a 2D matrix, where each matrix element holds the intensity value of each image pixel. This matrix corresponds to the model's input layer.

In the next stage, *feature learning*, the image is filtered using multiple *convolution operations* to detect diverse low-level features[4], e.g., edges, lines, and corners

[4]In this subsection, the term feature refers to the part or pattern of an image that helps its identification. The term is not to be confused with inspection features within the context of metrological tasks.

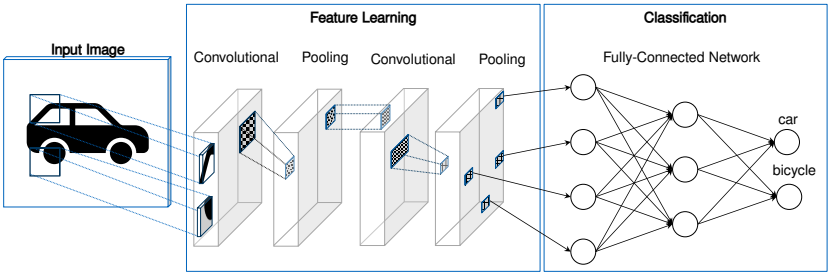


Figure A.4: Exemplary architecture and simplified outputs for the different layers of a CNN for image classification tasks based on BASLER (2021, p. 161) and MURPHY (2022, p. 472))

(GOODFELLOW et al. 2016, pp. 5–6). During this step, one kernel[5] is multiplied and shifted through the whole image to produce a *feature map* (CHOLLET 2018, pp. 122–123). In the following step, the feature map is downsampled using a pooling layer to obtain generalized and translation-invariant features (MURPHY 2022, p. 471). Generally, the architecture of CNNs comprises multiple pairs of convolutional and pooling layers to synthesize more complex features. The set of convolutional and pooling layers is known as the feature learning or detection component of CNNs (BASLER 2021, p. 161). The main task of this collection of layers is ultimately to convert image patches into numerical values so that the following layers of the neural network can extract relevant patterns from them.

In the last step, fully connected layers are used to interconnect the extracted features and ultimately detect, at a higher level of abstraction, a specific class of objects in an image. Due to their architecture, CNNs have found its prophecy to solve vision tasks efficiently, e.g., object identification, classification, and segmentation (BASLER 2021, p. 59).

A.1.3.3 Gaussian Processes

This subsection briefly summarizes the mathematical foundations and general characteristics of GPs.

Definition

A GP is a stochastic process, i.e., a collection of random variables over a temporal or spatial domain, in which each finite subset has a common Gaussian distribution (RASMUSSEN and WILLIAMS 2006, p. 13).

A GP is formally defined by its mean function $m(\mathbf{x})$ and a kernel denoted as $k(\mathbf{x}, \mathbf{x}')$ which models the covariance between two different input vectors $\boldsymbol{x}$

[5] In the context of image processing, kernels (related terms: filters, mask, convolution kernel) are matrices, mostly quadratic, used to filter information of an image.

and $\mathbf{x}'$:

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})], \tag{A.4}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\big[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))\big], \tag{A.5}$$

If a function $f(\boldsymbol{x})$ follows a GP, this is given as:

$$f(\mathbf{x}) \sim \mathcal{GP}\big(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\big), \tag{A.6}$$

where the symbol $\sim$ denotes that $f(\boldsymbol{x})$ follows a Gaussian distribution.

Kernel Design

In contrast to ANNs, GPs can be classified as memory-based procedures that make predictions based on similarity. Such procedures assume that observations with similar inputs will likely have similar target values. In the context of GPs, it is the kernel $k(\mathbf{x}, \mathbf{x}')$, which models the correlation between two inputs $\boldsymbol{x}$ and $\mathbf{x}'$. This property is considered one of the main strengths of kernel methods, as it allows the kernel to be designed based on prior knowledge and assumptions about the unknown function $f(\boldsymbol{x})$. For this reason, the choice or design of an appropriate kernel is of great importance, as it significantly determines how the random variables correlate with each other and ultimately affect the model's predictive performance. (BISHOP 2006, pp. 291–292; RASMUSSEN and WILLIAMS 2006, p. 79)

For example, the covariance between a pair of random variables $\mathbf{x}, \mathbf{x}' \in X$ can be computed using the *squared exponential kernel* (SE) (DUVENAUD 2014, p. 2):

$$\operatorname{cov}(f(\mathbf{x}), f(\mathbf{x}')) = k_{\text{SE}}(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\frac{(\mathbf{x} - \mathbf{x}')^2}{2\ell^2}\right), \tag{A.7}$$

where ℓ and σ represent hyperparameters to be learned during the training phase of the model. A list of other kernels is given by RASMUSSEN and WILLIAMS (2006, pp. 81–87).

Gaussian Process Regression

Assuming a historic dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ with n input and target value pairs $(\mathbf{x}_i, y_i)$, the following linear regression model is given

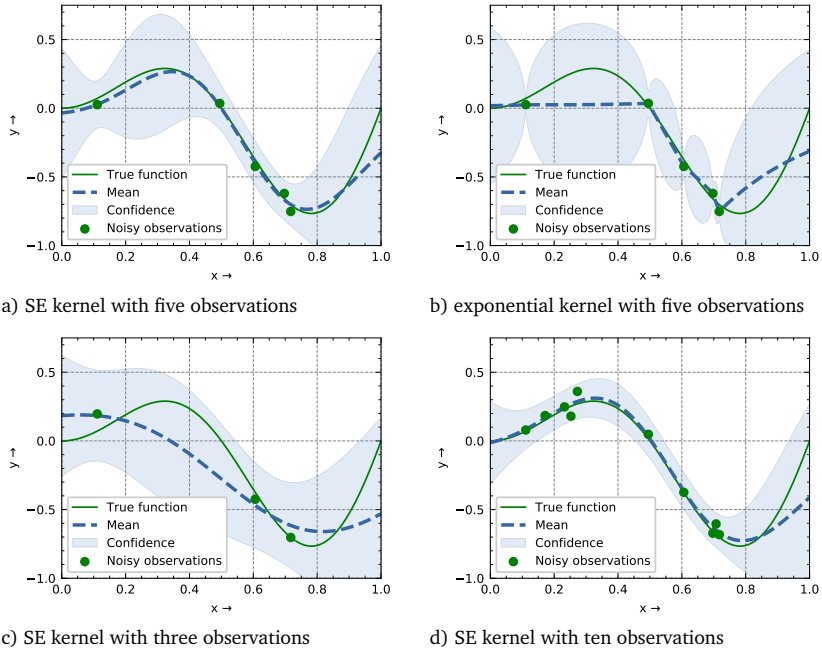
$$y_i = f(\mathbf{x}_i) + \epsilon, \quad \epsilon \sim \mathcal{N}\big(0, \sigma_\epsilon^2\big), \tag{A.8}$$

where ϵ denotes an additive independent distributed Gaussian noise with a zero mean and variance σ_ϵ^2.

The unknown function f is assumed to follow a GP. Hence, any subset of noisy observations $\mathbf{y} = (y_1, \ldots, y_n)^\top$ at the training input locations $X = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^\top$ follows a joint Gaussian distribution, which is expressed as follows:

$$\mathbf{y} \sim \mathcal{N}(0, \operatorname{cov}(\boldsymbol{y})), \text{ with}$$

$$\operatorname{cov}(\mathbf{y}) = K(X, X) + \sigma_n^2 I, \tag{A.9}$$



a) SE kernel with five observations

b) exponential kernel with five observations

c) SE kernel with three observations

d) SE kernel with ten observations

Figure A.5: Overview of different GPs, where the true function, $y = x\sin(2\pi x)$, (solid line in green) is reconstructed using noisy observations (dots in green) by a GP regression model (dashed line in blue) including its double standard deviation (blue area). The graphs on the top show GPs with different kernels and the graphs on the bottom with a different number of observations.

where I denotes a unity matrix of size $n \times n$ and $K(X,X)$ is a matrix of covariances. The covariance matrix elements are computed using a valid kernel[6], such as the SE kernel from Eq. A.7. Analogously, the prediction values for an unknown set of function values can be inferred based on the Bayes' Theorem for conditional distributions (see RASMUSSEN and WILLIAMS 2006, p. 16, pp. 200–201).

Fig. A.5 illustrates the key characteristics of GPs and significance of kernel design. An exemplary GP regression model, following a sinus function, is trained using two different kernels and a different number of observations. Fig. A.5a shows a SE kernel that has a smoother trend and a more accurate approximation than the exponential kernel in Fig. A.5b, although both kernels approximate all observations similarly well. Furthermore, Figures A.5c and A.5d demonstrate the influence of the number of training observations on prediction performance and confidence. The trends show that the integration of additional training points results in an improvement of both indicators.

[6]A valid kernel is defined to be symmetric if $k(\mathbf{x},\mathbf{x}') = k(\mathbf{x}',\mathbf{x})$. Furthermore, the covariance matrix must be positive semidefinite for all possible vectors, i.e., their eigenvalues are non-negative. (RASMUSSEN and WILLIAMS 2006, p. 80).

Moreover, the advantages of GP are visualized in all figures. First, all models show high flexibility to approximate nonlinearities, and useful predictions can be obtained with few data. Second, due to the stochastic nature of GPs, any prediction is automatically associated with an uncertainty. All trends in Fig. A.5 illustrate this behavior, showing the double standard deviation of the mean prediction and how the uncertainty increases within areas without training data. This characteristic is of great advantage in industrial applications, as it also allows for evaluating the prediction's reliability and the derivation of diverse actions.

A.2 Solution Modules

A.2.1 Requirements

Similar to the related research (CHEN and LI 2004; SCOTT 2002; TEKOUO MOUTCHIHO 2012), the following general requirements were considered in the present thesis as guidelines for the design and evaluation of the implemented technical solutions according to the ISO/IEC 25010 (2011):

1. Portability: The proposed approaches should be abstracted and generalized at the highest possible level. Model development should strive for a hardware-agnostic implementation, generalization to analogous applications, and the promotion of scalability.

2. Performance efficiency: The methods and techniques used should strive for a low algorithmic complexity to increase the overall computational efficiency of the solution modules.

3. Security: Due to traceability and safety issues within industrial applications deterministic approaches should be prioritized. Therefore, explicit, analytical, and linear models should be preferred whenever possible over heuristics.

4. Maintainability. The approaches and models should consider a modular structure and not affect other modules.

5. Usability: The parameters required to implement the models and approaches should be easily accessible to end-users. The proposed solution modules should strive a minimal parameterization.

The given order does not consider any prioritization of the requirements. In addition, these requirements are used to evaluate characteristics of the proposed solutions in Chapter 4.

A.3 Evaluation

A.3.1 Experimental Setup

This subsection introduces an overview of the hardware and software used to implement and evaluate the solution modules.

A.3.1.1 Reference ROMS

The industrial ROMS, *ZEISS AIBox* from the Carl Zeiss Optotechnik GmbH, was utilized as the reference system for implementing and evaluating all solution modules of the present thesis. The AIBox is an industrial robot-based 3D measurement cell comprising a SLS range sensor, a six-axis industrial robot, and a rotation table for fixing the inspecting workpiece. Fig. A.6 provides an overview of the core elements of the AIBox, which are further described below.

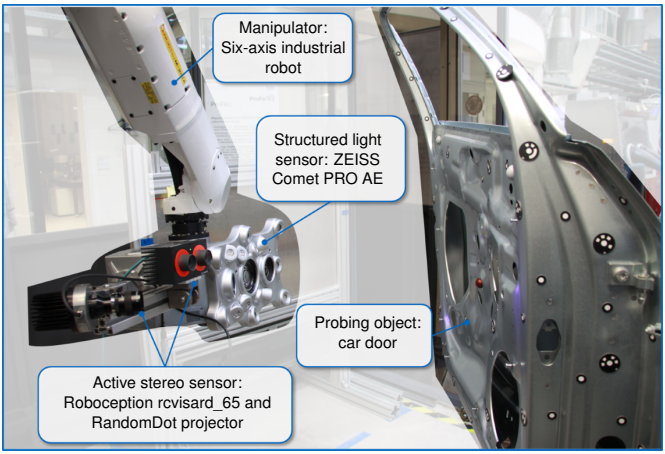


Figure A.6: Overview of the core components of the inspection AIBox extended by an active stereo system. (modified from MAGAÑA et al. (2023c))

A.3.1.2 Range Sensors

The *ZEISS Comet PRO AE* consists of a monochrome camera and a blue-light digital fringe projector. The camera includes a high-pass filter to minimize the impact of external lighting sources, allowing for the perception of blue light from the projector. In addition, the range sensor includes a photogrammetric camera that can be used for global image registration. The SLS can be used for different metrological tasks, e.g., coordinate metrology, digitization, and reverse engineering.

The sensor is controlled by the proprietary ZEISS *colin3D* software. In addition, the system integrates the proprietary software, ZEISS Visio7, which is used for the offline and online programming of the robot.

To evaluate the transferability of the solution modules with other range sensors, a stereo sensor (*rc_visard 65*) and a dot projector from *Roboception* were attached to the SLS.

A.3.1.3 Manipulators

A six-axis industrial robot (*Fanuc M20ia*) is used to position the sensor. Due to the limited workspace of the industrial robot, a rotary table is utilized to enable the rotation of workpieces and obtain measurements from all sides.

A.3.1.4 Workpiece and Features

To evaluate the applicability of the solutions presented in this thesis in an industrial context, a car door with 636 features was used as a probing object. The nominal position, orientation, and geometric dimensions of all features are known. In addition, the car door has different surface finishes on the inside and outside. The car door is mounted on the rotary table, but its relative pose to the robot base is initially unknown.

A.3.1.5 Software Architecture

All *solution modules* (SMs) were implemented as independent components using different programming languages and software frameworks. To integrate the developed modules with the proprietary software of the AIBox, different interfaces were developed. Fig. A.7 gives an overview of the reference architecture and the developed interfaces. A comprehensive overview of the implementation of the solution modules can be found in the corresponding publications.

A short overview of the most relevant software components is given below:

SM1: The synthetic pipeline was developed using the rendering engine from Unity (UNITY TECHNOLOGIES 2022) and a CAD model of the car door. In addition, the proposed CNN classification model for the orientation estimation was implemented in *Keras* (CHOLLET et al. 2015) and *TensorFlow* (TENSORFLOW DEVELOPERS 2023).

SM2: The backbone of the framework was developed using the *Robot Operating System* (ROS) (QUIGLEY et al. 2009). The framework was built using a knowledge-based service-oriented architecture design. The concept and architecture are described in MAGAÑA et al. (2019, 2020b). The offline computed viewpoint plans were transferred to the AIBox using an *extensible markup language* (XML) file defined by the proprietary software ZEISS Visio7.

SM4: The GP regression models were implemented in Python using the GPy library (GPy 2012). Moreover, two pipelines were implemented to automate the measurement acquisition and the post-processing of the acquired images and point clouds.

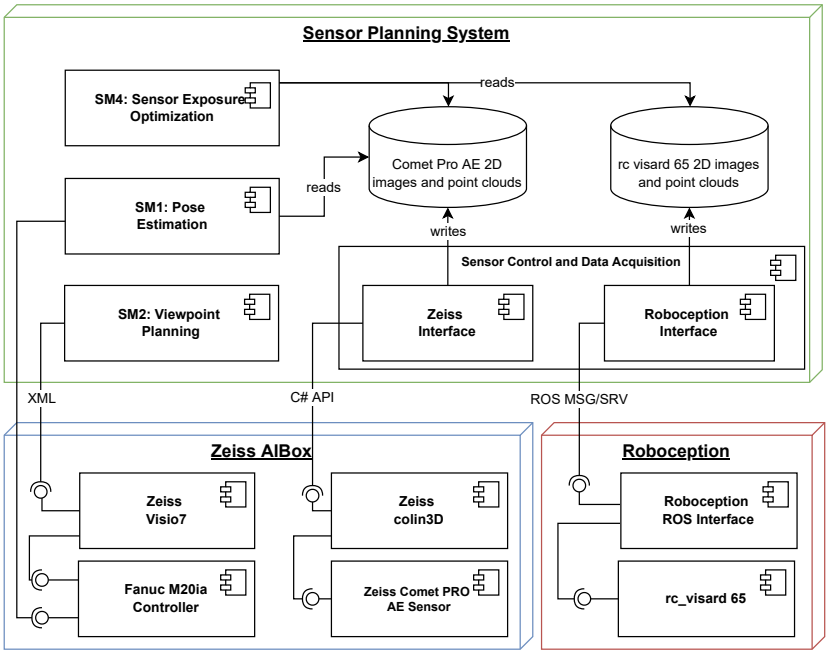


Figure A.7: Overview of the reference architecture and SMs. Top: Components developed within the scope of this thesis. Bottom left: Components of the Zeiss AIBox. Bottom right: Components of the Roboception rc_visard.

In addition, a handful of open source computer vision libraries were used through all solution modules, for camera calibration, image, and 3D mesh processing tasks (Bradski 2000; Dawson-Haggerty et al. 2022; Zhou et al. 2018; Zhou et al. 2016).

A.3.2 Economic Analysis

This section presents two statistical economic methods to analyze the economic benefits of the outlined solution of this thesis: a payback period analysis in Subsection A.3.2.2 and a break-even analysis in Subsection A.3.2.3. Before, a general cost analysis described in Subsection A.3.2.1 is introduced.

A.3.2.1 Costs Overview

In the first step, the investment costs and depreciation costs are estimated. Table A.1 provides an overview of these costs considering the following assumptions.

1. Solution Modules Software Development Costs ($c_{i,dev}$): Since the developed software tools are not publicly available, it is assumed that all software tools must be developed from scratch. The monetary cost for each module was calculated with an hourly wage rate[7] of $c_{op,h} = 60\,€/h$. The estimated development effort was calculated considering the expertise of a software engineer with a background in robotics, computer vision, and machine learning.

2. Initial Dataset Generation Costs ($c_{i,data}$): Taking into account that the ML models from PUB1 and PUB4 require an individual dataset for each workpiece, the effort for creating an initial dataset must be considered. Since PUB1 introduced a fully automated pipeline for generating synthetic datasets, a minimum effort of 10 h is considered to build an initial database consisting of different objects. On the contrary, an effort of 50 h was estimated to acquire an initial dataset that can be used to train the models proposed in PUB4.

3. Total investment (c_i): The total investment costs is the sum of all costs $c_i = c_{i,dev} + c_{i,data} + c_{i,other}$, where $c_{i,other}$ refers to additional fix costs, e.g., a workstation for model training, system integration, and planning costs.

4. Annual depreciation: A rough depreciation analysis is performed considering a six-year useful life[8]. In addition, assuming that all solutions are built on top of an existing production system and that the operating capacity over a year is the same, only the relative additional costs are considered (VOEGELE and SOMMER 2012, p. 346). For this reason, the operational machine costs are not included in the calculations.

5. Static methods: The analyses are carried out on the basis of static methods, which consider the costs and revenues to be constant over the period of use of an investment (VOEGELE and SOMMER 2012, p. 364).

A.3.2.2 Payback Period Analysis

To estimate the payback period of the proposed solutions, first the programming costs are calculated over one year for a reference system without them. Considering the reference programming effort $T_{ref} = 7.5\,\text{h}$ (see Eq. 4.1), this thesis assumes that a ROMS is capable of carrying out two reference inspection tasks

[7] The hourly wage rate was calculated from the employer's perspective considering an average monthly salary of a senior software engineer of 9600€ with a 40-hour week.

[8] The useful life is given in the depreciation table for the mechanical engineering industry (IV D 2-S 1551-470/01) of the German Federal Ministry of Finance.

per week and an average of eight per month. The yearly average programming effort, taking into account 96 inspection tasks per annum and an annual machine utilization of 80 % is equivalent to

$$p_{ref,a} = 0.8 \cdot 96/\text{a} \cdot 7.5\,\text{h} = 576\,\text{h/a}$$

with operator programming costs in the order of

$$c_{op,a} = c_{op,h} \cdot p_{ref,a} = 34.6\,\text{k€/a}.$$

The potential programming saving effort is then calculated based on the estimated technical efficiency evaluation, which corresponded to $p_{saving} = 87.6\,\%$ (see Eq. 4.2), as follows

$$p_{sav,a} = p_{saving} \cdot p_{ref,a} = 504.6\,\text{h/a}.$$

The annual operating cost of manual programming translates into a savings in the order of

$$c_{sav,a} = c_{op,h} \cdot p_{sav,a} = 30.3\,\text{k€/a}.$$

The expected payback period is calculated considering the total investment and yearly revenues of the system (VOEGELE and SOMMER 2012, p. 357). Taking into account the cost savings for manual programming, the depreciation of the system, and approximated operative costs of 4k€/a (e.g., workstation operation costs, technical support), the payback period is calculated as follows:

$$P = \frac{c_i}{c_{sav,a} + c_{dep,a} - c_{other,a}} = 2.2\,\text{a}$$

Under the assumptions considered in this subsection, it is concluded that the outlined approach is economically feasible after approximately two years.

A.3.2.3 Break-even Analysis

An additional economic evaluation driven by a cost comparison per feature is introduced to shed light on the potential of the approach from another perspective, i.e. rationalization investments. This type of investment evaluates whether an existing system should be replaced by a new approach in order to reduce costs or improve quality (LÜCK 2004, p. 171; VOEGELE and SOMMER 2012, p. 345).

Assuming a linear programming effort per feature and the programming planning effort T_{ref}, the manual operating cost for programming a feature (f) for the reference scenario is

$$c^f_{ref} = T_{ref} \cdot \frac{c_{op,h}}{500\,f} = 7.5\,h \cdot \frac{60\,€/h}{500\,f} = 0.9\,€/f$$

and for the approach of this thesis:

$$c^f_{semi-aut} = (1 - p_{saving}) \cdot c^f_{ref} = 0.11\,€/f\,.$$

The break-even point is calculated using a static analysis of the incremental costs of the proposed solutions, which corresponds to the total investment. In the most conservative scenario, it is anticipated that there will be no revenue. Moreover, it is assumed that the depreciation costs should cover the annual expenses of the proposed solutions, which are given for the assumed numbers in Tab. A.1, i.e., $c_{dep,a} > c_{other,a}$. The break-even number of features is calculated with the given formula (VOEGELE and SOMMER 2012, pp. 347–348):

$$n_{break} = \frac{c_i}{c^f_{ref} - c^f_{semi-aut}} = 118,228\,f.$$

According to the break-even analysis, an economic benefit is expected after programming the inspection of 118,228 features or 237 inspection orders with the complexity of a car door with 500 features.

Table A.1: Overview of costs for economic evaluation

	Investment	
	Software Development per solution module (SM)	
	SM1: Dataset Pipeline (160 h)	9.6 k€
	SM1: ML Models (160 h)	9.6 k€
	SM2: Viewpoint Planning Framework (500 h)	30 k€
	SM3: Dataset Pipeline (160 h)	9.6 k€
	SM3: ML Models (160 h)	9.6 k€
$c_{i,dev}$	$\sum$ Total (1140 h)	68.4 k€
	Dataset Generation	
	SM1: Dataset generation (10 h)	0.5 k€
	SM3: Dataset generation (50 h)	3.0 k€
$c_{i,data}$	$\sum$ Total (60 h)	3.5 k€
$c_{i,other}$	Other	25.0 k€
c_i	$\sum$ Total investment	93.4 k€
	Annual Depreciation	
	Utilization period	6 a
$c_{dep,a}$	Linear depreciation	15.6 k€/a
	Programming Costs	
$c_{op,h}$	Operator programming cost per hour	60 €/h
	Annual costs reference metrology task	
$p_{ref,a}$	Programming hours per year	576 h/a
$c_{op,a}$	Operator programming costs per year	34.6 k€/a
	Annual costs reference metrology task with semi-automatized programming	
p_{saving}	Saving programming effort factor	87.6 %
$c_{sav,a}$	Saved operator programming costs per year	30.3 k€/a
$c_{other,a}$	Operative costs per year	4 k€/a
	Cost-benefits	
P	Payback period	2.2 a
n_{break}	Break-even number of inspecting features	118,228

Appendix B

Research and Academic Activities

B.1 List of Publications

Within the framework of his scientific activities at the *Institute for Machine Tools and Industrial Management of TU Munich (German: Institut für Werkzeugmaschinen und Betriebswissenschaften der TU München)* (*iwb*), the author collaborated on the following publications

Peer Reviewed

1. MAGAÑA and REINHART (2017): „Herstellerneutrale Programmierung von Robotern“
2. IMHOFF, SCHNELL, MAGAÑA, DIERMEIER, SCHEIDERER, BRAUN, IMHOFF, ARCIERO, and BEITZEL (2018): „Single cut distal femoral osteotomy for correction of femoral torsion and valgus malformity in patellofemoral malalignment - proof of application of new trigonometrical calculations and 3D-printed cutting guides“
3. MAGAÑA, BAUER, and REINHART (2019): „Concept of a learning knowledge-based system for programming industrial robots“
4. BAUER, MAGAÑA, and REINHART (2019b): „Free-Form Surface Analysis and Linking Strategies for High Registration Accuracy in Quality Assurance Applications“
5. BAUER, GONNERMANN, MAGAÑA, and REINHART (2019a): „Autonome Prüfsysteme in der digitalen Fabrik: Skill-basierte Modellierung in der geometrischen Qualitätsprüfung für roboterbasierte Messsysteme“
6. BAUER, FINK, MAGAÑA, and REINHART (2020b): „Spatial interactive projections in robot-based inspection systems“
7. BAUER, LI, MAGAÑA FLORES, and REINHART (2020a): „Accuracy Analysis of Alignment Methods based on Reference Features for Robot-Based Optical Inspection Systems“

8. MAGAÑA, GEBEL, BAUER, and REINHART (2020b): „Knowledge-Based Service-Oriented System for the Automated Programming of Robot-Based Inspection Systems“

9. MAGAÑA, WU, BAUER, and REINHART (2020a): „PoseNetwork: Pipeline for the Automated Generation of Synthetic Training Data and CNN for Object Detection, Segmentation, and Orientation Estimation“

10. BAUER, HECKLER, WORACK, MAGAÑA, and REINHART (2021b): „Registration strategy of point clouds based on region-specific projections and virtual structures for robot-based inspection systems“

11. BAUER, GOTTSCHALL, MAGAÑA, HOFER, and REINHART (2021a): „A Technological and Economic Potential Analysis of Measurement Systems in Geometrical Quality Assurance“

12. BAUER, SCHMITT, DIRR, MAGAÑA, and REINHART (2022): „Intelligent predetection of projected reference markers for robot-based inspection systems“

13. MAGAÑA, DIRR, BAUER, and REINHART (2023a): „Viewpoint Generation Using Feature-Based Constrained Spaces for Robot Vision Systems“

14. MAGAÑA, VLAEYEN, HAITJEMA, BAUER, SCHMUCKER, and REINHART (2023b): „Viewpoint Planning for Range Sensors Using Feature Cluster Constrained Spaces for Robot Vision Systems“

15. MAGAÑA, SCHNEIDER, BENKER, ALTMANN, BAUER, and REINHART (2023c): „Exposure Time and Point Cloud Quality Prediction for Active 3D Imaging Sensors using Gaussian Processes“

Preprint

16. MAGAÑA, DIRR, BAUER, and REINHART (2023d): „Viewpoint Generation using Feature-Based Constrained Spaces for Robot Vision Systems“

Book

17. REINHART, MAGAÑA FLORES, and ZWICKER (2018): *Industrieroboter. Planung - Integration - Trends: ein Leitfaden für KMU*

Patents

18. TSIOUVARAS, DAUB, BACHMANN, SEMM, RICHTER, MAGAÑA, and METKAR (2019): „Prüfvorrichtung und Verfahren zum Testen einer Speicherzelle“

19. BAUER and MAGAÑA (2019): „Verfahren und System zur dreidimensionalen Erfassung einer Szene“

B.2 Supervised Student Theses

Within the framework of this research, the student theses listed below were carried out at the *iwb* of the *Technical University of Munich* (TUM) under the scientific, technical and content-related supervision of the author of this dissertation. In these projects, various research questions concerning the improvement of the individual *solution module* (SM) for the automation of *robot-based optical 3D measuring system* (ROMS) were investigated. Some of the results have been incorporated into the publications related to the present work. The author would like to take this opportunity to express his sincere thanks to all students for their commitment in supporting this scientific work.

SM1

- Farrand, Kevin. *Automatic Recognition and Alignment of Measurement Object based on a Robot-based Optical Inspection System.* Master Thesis. 2018
- Hang, Wu. *Workpiece detection for an autonomous robot-based measurement system based on Machine Learning.* Master Thesis. 2019

SM2

- Gebel, Silvia. *Concept for Automated Pose Planning for Hardware-neutral Optical Robotic Measurement Systems.* Bachelor Thesis. 2018
- Longares Díez, Joaquín María. *Sensor planning optimization for robot-based optical inspection with visibility constraints and overlap control for registration*. Master Thesis. 2019
- Pritscher, Stefan Anton. *Aufnahmeposen-Optimierung für roboterbasierte Messsysteme durch Schnittvolumensuche zwischen Lösungsräumen.* Semester Thesis. 2019
- Schäfftlein, Ansgar. *Messposenoptimierung für ein roboterbasiertes Messsystem.* Semester Thesis. 2019
- Xiaoxuan, Ye. *Sensor Planning Optimization with Registration Control for RobotBased Optical Inspection Systems.* Semester Thesis. 2021

SM3

- Treis, Marcel. *Supervised Machine Learning zur Klassifizierung von Aufnahmeposen und -parametern für ein cyberphysisches, roboterbasiertes optisches 3D-Messsystem.* Master Thesis. 2019

- Altmann, Thomas. *Machine Learning zur automatischen Optimierung von Aufnahmeposen für ein cyberphysisches, roboterbasiertes optisches 3D-Messsystem*. Master Thesis. 2019
- López López, Adriana. *Gauss Process Models for the Automated Optimization of Measurement Parameters of a Robot-based 3D Inspection System.* Master Thesis. 2020
- Flehmig, Niclas. *Automatisierte und optimierte Parametrisierung von 3D-Messposen für roboterbasierte optische Messsysteme auf Basis von Gaußprozessen.* Semester Thesis. 2022
- Qiu, Stefan. *Modellerweiterung und Evaluation von Gauß-Prozessen Modellen für die automatische Optimierung von Messparametern für ein robotergestütztes optisches 3D-Messsystem.* Semester Thesis. 2021
- Schneider, Lukas. *Automatisierung eines robotergestützten StereoKamera-Messsystems auf Basis von Gauß- Prozessen und maschinellem Sehen.* Master Thesis. 2022

Additionally, cross-cutting issues that contributed to the implementation of the foundations of this dissertation were investigated in the following theses:

- Stoiber, Manuel. *Development Environment for the Research of Smart Robots.* Bachelor Thesis. 2017
- Roßmann, Jonas. *Herstellerneutrale Modellierung und Inbetriebnahme von Industrierobotern*. Bachelor Thesis. 2017
- Abdelaziz Abdelsattar Abdelaziz, Mohamed. *Vergleich von Bahnplanungsalgorithmen für ein roboterbasiertes Messsystem*. Bachelor Thesis. 2017
- Rupert, Tobias. *Aufbau eines Schnittstellenmodells für ein wissensbasiertes System zur autonomen Programmierung eines roboterbasierten optischen Messsystems.* Semester Thesis. 2019
- Jaeho, Cha. *Softwaretool zur automatischen Erstellung von kinematischen Modellen für Robotersysteme*. Bachelorarbeit. 2021
- Theobald, Marc. *Implementierung eines wissensbasierten und serviceorientierten Systems für die automatisierte Programmierung von robotergestützten Inspektionssystemen.* Semester Thesis. 2021