

The Working Posture Controller: Automated adaptation of the workpiece pose to enable a neutral working posture

vorgelegt von
Dipl.-Inf. The Duy Nguyen
geb. in Ho-Chi-Minh-Stadt

von
der Fakultät V - Verkehrs- und Maschinensysteme
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
- Dr.-Ing. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr.-Ing. Rainer Stark
Gutachter:	Prof. Dr.-Ing. Jörg Krüger
Gutachter:	Prof. Dr.-Ing. Jochen Deuse

Tag der wissenschaftlichen Aussprache 18.10.2017

Berlin, 2018

Vorwort des Autors

Diese Arbeit entstand am Fachgebiet Automatisierungstechnik am Institut für Werkzeugmaschinen und Fabrikbetrieb (IWF) der TU Berlin im Rahmen meiner Arbeit am Forschungsprojekt "SFB 1026 - Sustainable Manufacturing".

Bevor der wissenschaftliche Teil beginnt, möchte ich die Gelegenheit nutzen einigen Menschen zu danken.

Ein großer Dank geht an meinen Betreuer, *Prof. Dr.-Ing. Jörg Krüger*, für sein Interesse, seine großzügige Förderung und seine Unterstützung während der fünf Jahre.

Außerdem danke ich *Prof. Dr.-Ing. Jochen Deuse* vom Institut für Produktionssysteme (IPS) der Technischen Universität Dortmund für die Übernahme der Gutachtertätigkeit und dem Vorsitzenden des Promotionsausschusses, *Prof. Dr.-Ing. Rainer Stark*.

Ich möchte mich auch bei den zahlreichen Kollegen des Fachgebietes, des Fraunhofer IPK und des Sonderforschungsbereichs bedanken, die für ein sehr angenehmes Arbeitsklima sorgten, in dem ich viel lernen konnte. Insbesondere möchte ich *Ekaterina Ivanova*, *Martin Rudorfer*, *Bernhard Föllmer*, *Christian Bloch* und *Gregor Thiele* erwähnen, die Teile des Manuskripts gelesen und wertvolle Hinweise gegeben haben. Ohne die Unterstützung der Kollegen *Hanna Lickert*, *Bernd Muschard*, *Udo Templiner* und *Arne Glodde*, hätte ich mein Erdachtes nicht in etwas Anfassbares umsetzen können. Mein Dank gilt auch *Martin Kleinsorge*, der mir geholfen hat mich in den wissenschaftlichen Alltag einzufinden.

Ich danke auch meinen ehemaligen Studenten *Carla Pilz*, *Arne Fey*, *Tim Sziburis*, *Adrian Hill* und *Max Baumgarten*, die wichtige Unterstützung bei der Implementierung geleistet haben.

Durch meine Freunde *Doan Huong Giang*, *Chu Hoang Tung* und *Farhad Arbabzadah* konnte ich für ein paar Stunden den Stress der Dissertation vergessen.

Ein besonderer Dank geht an *Vera Scholwin*, die mich fachlich und moralisch unterstützt hat und viel Geduld und Verständnis aufbrachte.

Zum Schluss möchte ich meinen Eltern *Nguyen Thanh Nguyen* und *Nguyen The Tuyen* danken, die mich zu diesem Weg ermutigt haben und immer an mich glaubten.

Berlin, Februar 2018

The Duy Nguyen

Contents

Abbreviations and Symbols	xiv
Kurzfassung	xvii
Abstract	1
1 Introduction	2
1.1 Work-related Musculoskeletal Disorders	2
1.2 Contribution of this Thesis	4
1.3 Thesis Outline	4
2 State of the art	6
2.1 Terminology	6
2.2 Ergonomic Assessment	7
2.2.1 Self-Reports	9
2.2.2 Direct Methods	9
2.2.3 Observational Methods	10
2.2.4 Summary of Ergonomic Assessment Methods	12
2.3 Ergonomic Intervention	14
2.3.1 Interventions against Force Exposure	15
2.3.2 Interventions against Repetition	15
2.3.3 Interventions against Awkward Posture	16
2.3.4 Computer-Aided Ergonomic Intervention	16
2.3.5 Impact of Ergonomic Interventions	23
2.3.6 Summary of Ergonomic Intervention Methods	25
2.4 Conclusions	27
3 Research Goal	28
3.1 Research Gap	28
3.2 Research Hypotheses	29
3.3 Research Approach	31
4 The Working Posture Controller	33

4.1	Problem Statement	33
4.2	Design Alternatives	34
4.2.1	Detecting an Awkward Posture	34
4.2.2	Measures against Postural Load	36
4.3	System Concept of the Working Posture Controller (WPC)	38
4.4	Use Cases	38
4.5	Safety Mechanisms	40
4.6	Conclusions	43
5	Posture Assessment	44
5.1	Problem Statement	44
5.2	Design Alternatives	45
5.2.1	Posture Classification	45
5.2.2	Joint Localisation	50
5.2.3	Image Acquisition Devices	55
5.3	Posture Classification Workflow	59
5.4	Camera Setup	60
5.5	Joint Localisation	63
5.5.1	Pre-processing of the Image Data	64
5.5.2	Notation	65
5.5.3	2D Joint Localisation	66
5.5.4	3D Joint Localisation	73
5.6	Posture Classification	78
5.7	Conclusions	80
6	Posture optimisation	81
6.1	Problem Statement	81
6.2	Design Alternatives	82
6.2.1	Posture Prediction	82
6.2.2	Workpiece Pose Search and Workplace Design	84
6.2.3	Control Strategy	86
6.3	Optimisation-Based Approach	88
6.4	Optimisation Objectives	90
6.4.1	Posture Objective	91
6.4.2	Centroid Objective	92
6.4.3	Working Direction Objective	92
6.4.4	Contact Point Objective	93
6.5	Modelling for Height-Adjustable Workplaces	93
6.6	Modelling for Robot Kinematics	95
6.7	Solving the Optimisation Problem	98
6.8	Control Strategies	99
6.8.1	Temporal EAWS Threshold (TET) Control Strategy	100

6.8.2	Decaying Temporal EAWS Threshold (DTET) Control Strategy . . .	100
6.9	Conclusions	103
7	Implementation	105
7.1	Scenario	105
7.2	Hardware	106
7.2.1	Sensor and Input Device	106
7.2.2	Actuators	107
7.3	Software	108
7.3.1	WPC System Architecture	108
7.3.2	The WPCApp	110
8	Evaluation	113
8.1	Experimental Datasets	114
8.2	Processing Time	117
8.3	Posture Assessment	119
8.3.1	Standing Postures (DAT1)	119
8.3.2	Comparison to Alaska/EAWS Assessment (DAT2)	124
8.4	Posture Optimisation	127
8.4.1	Comparison of Height-Adjustable Actuator and Robot Actuator (DAT1)	127
8.4.2	Comparison of TET and DTET Control Strategies (DAT3)	132
8.5	Conclusions	137
9	Conclusions and Outlook	139
9.1	Summary	139
9.2	Achievements of this Thesis	140
9.3	Limitations of this Thesis	142
9.3.1	Focus on Posture	142
9.3.2	3D Joint Localisation Algorithms	142
9.3.3	Mathematical Modelling of Human Behaviour	143
9.3.4	Financial Amortisation	143
9.3.5	Ethical, Social, Legal implications (ELSI)	144
9.4	Outlook	144
A	Posture Assessment using EAWS	146
B	DIN 33402: Mean Lengths of Body Parts in cm	151
C	The SQP Framework	152
D	Implementation Details	154

D.0.1	Libraries and Tools	154
D.0.2	Division of Components into Packages	155
D.0.3	Performance Optimisations through Parallel Computing	156
E	Supplementary Experimental Results	158
	Bibliography	161

List of Figures

1.1	Reasons for worker absenteeism in Germany in 2009 according to the German Public Health Fund (AOK) ordered by prevalence (adapted from Badura et al. [1]).	3
2.1	Physical exposure dimensions (adapted from Winkel and Mathiassen [2]). .	7
2.2	Classification of work-related musculoskeletal disorder (WMSD) risk factors and the focus of the current work.	8
2.3	Overview of ergonomic assessment methods according to Li and Buckle [3] (second row) and Ellegast [4] (third row). The roles in the arrow denote the user groups the tools are targeting.	8
2.4	Examples of direct methods.	10
2.5	Summary of ergonomic assessment methods (adapted from Winkel and Mathiassen [2]).	13
2.6	Classification of ergonomic interventions according to Goggins et al. [5]. . .	15
2.7	Comparison of traditional workplace development costs with Digital Human Model (DHM) workplace development costs (adapted from Delleman et al. [7]).	18
2.8	Visual feedback for working postures proposed by Vignais et al. [9]. . . .	22
2.9	Illustration of Cobots and exo-skeletons.	23
2.10	Safety and ergonomics simulation of a robot cell described in Thomas et al. [10].	24
4.1	Summary of the sub-research questions RQ1 implies.	34
4.2	Measures for avoiding awkward postures.	36
4.3	An actuator holding the workpiece is used to adjust the workplace. The left panel illustrates an ergonomically unfavourable actuator configuration, which is corrected in the right panel.	37
4.4	The basic WPC control loop.	39

4.5	Summary of safety mechanisms and the threats they address.	41
4.6	Safety enhanced WPC workflow.	42
5.1	Summary of sub-research questions RQ2 implies.	45
5.2	Overview of feature types used for posture classification.	46
5.3	Examples of direct features for posture classification.	47
5.4	Ambiguity of joint angle features. Both postures are represented by the same joint angle vector as the joint angle values q_1 and q_2 are similar.	49
5.5	Shotton et al. [11] used intermediate body part maps to compute joint coordinates.	51
5.6	Shortcomings of the joint localisation algorithm of Shotton et al. [11] and their implementation in the Primesense NiTE library. The middle picture suggests that the posture adopted was not present in the training data set and, thus, cannot be recognised.	52
5.7	Colour image (left) and corresponding depth image (right) from the Microsoft Kinect® sensor.	56
5.8	Illustration of the developed posture classification workflow.	59
5.9	Camera coordinate system notation used in this thesis.	60
5.10	Visualisation of variables influencing the required spatial resolution.	61
5.11	Illustration of the variables z_{min} , α_{FOV} and a	62
5.12	Intermediate results of the pre-processing pipeline.	64
5.13	Illustration of the pre-processing workflow.	65
5.14	Human model terminology used in the following sections.	65
5.15	Workflow of the 2D joint localisation approach.	66
5.16	Visualisation of the 2D human model.	67
5.17	Silhouette image and overlaid model triangles.	68
5.18	Visualisation of the graph built from a solution set with $J = 2$ and $N = 2$. An exemplary path from v_s to v_t is coloured in green.	71
5.19	Intermediate results while preparing the fitting of the upper limb model. . .	71
5.20	The problem of the shortened upper limb and the most successful solution. .	73
5.21	Visualisation of the ray-casting approach.	74

5.22	Exemplary result of the 3D joint localisation algorithm.	78
5.23	Flow diagram of 2D and 3D joint localisation approach.	79
6.1	Input and output of the posture optimisation algorithm for the example of a height-adjustable actuator.	82
6.2	Summary of sub-research questions RQ3 implies.	83
6.3	Mainprice et al. [13] used an optimisation based approach, however, the objective function does not consider the posture of the subject.	85
6.4	Methods engineering workflow.	85
6.5	Notations used for describing the details of the posture optimisation approach.	90
6.6	Visualisation of the "ideal" posture, \mathbf{q}_{ideal} , defined in the EAWS.	91
6.7	Visualisation of the centroid objective.	92
6.8	Robot model notation used to describe the proposed approach.	96
6.9	Visualisation of the variables used in the optimisation problem for the robot actuator.	97
6.10	Visualisation of Ergonomics Assessment Worksheet (EAWS) score (black lines denote EAWS scores 25 and 50) and height-adjustable actuator offset over time for the TET strategy.	101
6.11	Visualisation of EAWS and height-adjustable actuator offset over time for the DTET strategy.	103
6.12	Control strategy integrated into the WPC loop.	104
7.1	The WPC demonstrator scenario.	106
7.2	Motion range of the bicycle frame holder.	107
7.3	Robot actuator used in the demonstrator scenario.	108
7.4	Package dependencies within the WPC.	109
7.5	The WPCApp graphical user interface.	111
7.6	Visualisation of the Model-View-Controller (MVC) pattern.	111
7.7	Observer pattern applied when the model updates the view.	112
8.1	Research questions to be answered in the evaluation experiments.	113
8.2	Sample frames from DAT1.	116

8.3	Sample frames from DAT2.	117
8.4	Workflow to obtain EAWS postures from subject motions.	117
8.5	One exemplary posture sequence from DAT1 converted into a DAT3 sequence.	118
8.6	Visualisation of the feature space for different approaches using t-SNE. Different colours refer to varying class labels.	121
8.7	Confusion matrices achieved with joint angle features from different localisers.	122
8.8	EAWS score comparison for the 2D + RBF-SVM combination.	123
8.9	Difference between 2D+RBF-SVM labelling and manual labelling.	125
8.10	Difference between 2D+RBF-SVM labelling and Alaska/EAWS labelling.	126
8.11	Difference between manual labelling and Alaska/EAWS labelling.	127
8.12	Qualitative posture optimisation results.	128
8.13	Distribution of postures before and after optimisation for height-adjustable and robot actuator.	129
8.14	Distribution of working direction angles in DAT1 before posture optimisation.	130
8.15	Mean absolute difference of working directions as a function of the original working direction value.	130
8.16	Distribution of working direction values before and after optimisation. The lower the transparency of the data points the higher the density.	131
8.17	Interaction between the testbench components.	133
8.18	Distribution of non-dominated points.	135
A.1	Posture section of the EAWS.	148
A.2	Example of how to calculate an ergonomic risk score.	149
A.3	Exemplary case where a higher EAWS score does not represent a higher share of high-load postures.	150
A.4	Assymetry effects (3D postures) defined in the EAWS.	150
D.1	Parallelisation of the tree search. One CPU core processes the green nodes while the other processes the gray nodes.	157

E.1	Influence of strategy parameters on EAWS and ADAP.	158
E.2	Influence of actuator range on EAWS and ADAP.	159
E.3	Influence of posture duration on EAWS and ADAP.	160

List of Tables

2.1	Summary of observational methods.	11
2.2	Interventions classified according to effectiveness and addressed risk factor.	26
2.3	Interventions assigned to either pre-process or in-process measures.	26
2.4	Maximum achievable effectiveness level for different risk factors. The cell colours correspond to the effectiveness levels depicted in Fig. 2.6.	27
3.1	Maximum achievable level of effectiveness after successful accomplishment of the research goal.	30
4.1	Overview of awkward posture detection alternatives. Columns highlighted in green represent the selected approach.	35
4.2	Overview of awkward posture avoidance alternatives. Columns highlighted in green represent the selected approach.	37
5.1	Overview of alternative features for posture classification. Columns highlighted in green represent the selected approach.	50
5.2	Overview of joint localisation alternatives. Columns highlighted in green represent the selected approach.	54
5.3	Overview of image acquisition device alternatives. Columns highlighted in green represent the selected approach.	59
6.1	Overview of posture prediction alternatives. Columns highlighted in green represent the selected approach.	84
6.2	Overview of workpiece pose search alternatives. Columns highlighted in green represent the selected approach.	86
8.1	Summary of datasets (number of sequences and content) and their usage in the experiments.	114
8.2	Mean accuracy and mean EAWS differences for the tested feature-classifier combinations. Best values are marked bold.	120

8.3	Number of non-dominated TET and DTET data points for various posture length and standard height strategies.	135
A.1	Angle intervals of posture classes (in degrees) for joints in EAWS and REBA.	146
B.1	Lengths of parts of the body (fifth percentile) according DIN 33402.	151

Abbreviations and Symbols

Vectors are denoted by bold lower case symbols, e. g. \mathbf{a} .

Matrices are denoted by bold upper case symbols, e. g. \mathbf{A} .

Specifiers are denoted in the subscript. e. g. \mathbf{x}_{obj} .

Vector or matrix components are denoted in a subscript, as well. e. g. $\mathbf{x}_{obj,y}$.

Abbreviations

<i>ADAP</i>	accumulated distance of the actuator path
<i>Cobot</i>	collaborative robot
<i>DHM</i>	digital human model
<i>DoF</i>	degree of freedom
<i>DTET</i>	decaying temporal EAWS threshold
<i>EAWS</i>	ergonomic assessment worksheet
<i>EMG</i>	electromyography
<i>FoV</i>	field of view
<i>GUI</i>	graphical user interface
<i>IK</i>	inverse kinematics
<i>IMU</i>	inertial measurement unit
<i>KNN</i>	K-nearest neighbour
<i>MOO</i>	multiple objective optimisation
<i>MSD</i>	musculoskeletal disorder
<i>MTM</i>	methods time measurement
<i>MVC</i>	model-view-controller
<i>PSO</i>	particle swarm optimisation
<i>QP</i>	quadratic program
<i>RF</i>	random forest
<i>RGB</i>	red green blue
<i>RGBD</i>	red green blue - depth
<i>SQP</i>	sequential quadratic programming
<i>SVM</i>	support vector machine
<i>TET</i>	temporal EAWS threshold
<i>TOF</i>	time-of-flight
<i>WMSD</i>	work-related musculoskeletal disorder

WPC working posture controller

Symbols

Image Acquisition Devices (Sec. 5.2.3 - 5.4)

a	Minimal extent of viewable space (fig. 5.11)
α_{FOV}	Minimal field of view angle (fig. 5.11)
α_{min}	Minimal size of all angular intervals
B	Baseline length between two cameras
dD	Achievable disparity resolution
dZ	Size of a depth element
Δ_{min}	Minimal required spatial resolution (fig. 5.10b)
f	Focal length of both cameras
l	Body segment length
Z	Distance of a 3D point to the camera centre
z_{min}	Minimal distance of camera to scene (fig. 5.11)

Joint Localisation (Sec. 5.5 until end of Chapter 5)

c	Term weight
\mathbf{c}	Centroid of ellipsoid
\mathbf{e}	Support vector of a ray
d	Distance function
\mathbf{d}	Direction vector of a ray
h	Height of a segment
I	Image
$I_{\mathcal{M}}$	Model image
I_S	Model image of one isolated segment
J	Number of joints
\mathcal{M}	Kinematic Model
N	Number of discrete joint samples
\mathbf{q}	Joint angle vector
\mathbf{q}_{global}	Globally best particle
\mathbf{q}_{local}	Best position of a particular particle so far
\mathbf{R}	Rotation matrix
\mathcal{S}	Segment of kinematic model
\mathbf{S}	Scaling matrix

t	Ray parameter
v	Vertex in a graph
w	Width of a segment
\mathbf{x}	Position vector

Posture Optimisation (Sec. 6.3 - 6.7)

c	Term weight
\mathbf{d}	Direction vector
\mathbf{d}_w	Working direction vector
f	Objective function
h	Height of a segment
J	Number of joints
$l_{Contact}$	Length of vector from contact point to object front point
\mathbf{p}	Actuator parameter vector
Φ	Global angle representation of 2D vector
\mathbf{q}	Joint angle vector
\mathbf{x}	Position vector

Control Strategy (Sec. 6.8 until end of Chapter 6)

LUT	Function represented as lookup-table
n	Number of frames
P	Posture sequence
S_{EAWS}	EAWS score
t	Current time or frame
T	Threshold value
τ	Time share

Evaluation (Chapter 8)

acc	Accuracy
\mathbf{C}	Confusion matrix
n	Number of frames
P	Posture sequence
S_{EAWS}	EAWS score

Kurzfassung

Der Mensch hat eine prägende Rolle in der Produktion. Moderne Automatisierungstechnik hat schon große Teile der Belegschaft in Fabriken ersetzt. Jedoch gibt es Szenarien, wo eine Vollautomatisierung nicht realisiert werden kann. Zum Beispiel sind die hohen Rekonfigurationskosten der sich ständig ändernden Produktionslinien für Unternehmen, die ein breites Spektrum an Produkten anbieten, oft wirtschaftlich schwer darstellbar. Des Weiteren gibt es immer noch Aufgaben, wie das Handhaben flexibler Bauteile, die noch nicht effektiv automatisiert werden können. Erfahrene Werker sind im Gegensatz zur Maschine in der Lage, neue Aufgaben in kurzer Zeit zu erlernen. Ausserdem besitzen sie kognitive und sensormotorische Fähigkeiten, die derzeit von keiner Maschine kopiert werden können. Andererseits stellt der Erhalt der Werker Gesundheit eine Herausforderung dar. Arbeitsbezogene Muskel- und Skelletterkrankungen sind die vorwiegende Ursache von Fehltagen bei Werkern. Sie führen zu hohen finanziellen Verlusten bei Staat und Unternehmen. Eine wesentliche Ursache dieser Erkrankungen ist, dass die Tätigkeiten das Einnehmen von Zwangshaltungen erfordern.

Aufgrund ihrer großen Bedeutung standen Zwangshaltungen und deren Vermeidung bereits in der Vergangenheit im wissenschaftlichen Fokus. Bisher vorgestellte Lösungen führten zu einer signifikanten Verbesserung der Situation. Jedoch ist die Umsetzung der effektivsten Ansätze aufwändig und schränkt die Flexibilität der Fertigungslinie ein. Mit anderen Worten: Effektive Lösungen müssen an Tätigkeit und Arbeitsplatz angepasst werden. Zwar gibt es Lösungsansätze, die sich automatisch an neue Tätigkeiten anpassen können. Diese sind jedoch wesentlich ineffektiver. In der vorliegenden Arbeit wird dieses geschilderte Dilemma als "effectiveness-flexibility trade-off" bezeichnet.

Diese Dissertation versucht den oben beschriebenen Zielkonflikt zu überbrücken. Es wurde der Working Posture Controller (WPC) vorgestellt. Dieses neuartige Arbeitsmittel imitiert die Tätigkeit von Arbeitsplatzgestaltern. Der WPC bewertet automatisch die Körperhaltung des Werkers während der Ausführung der Arbeit. Tritt eine ergonomisch ungünstige Situation auf, schlägt der WPC eine Umgestaltung des Arbeitsplatzes vor, die dem Werker eine Ausführung der gleichen Tätigkeiten in einer ergonomisch besseren Haltung ermöglicht. Der eigentliche Umgestaltungsprozess kann dann mittels einer Aktorik realisiert werden. In der vorliegenden Arbeit wurden Systemkonzept und Basisalgorithmen zur Realisierung eines solchen Systems entwickelt. Im Detail wurden Methoden vorgeschlagen, um automatisch sowohl ergonomische Haltungsbewertungen mit einer Tiefenkamera durchzuführen als auch diese Haltung durch Umgestaltung des Arbeitsplatzes zu optimieren. Da die Rechenzeit nur wenige Sekunden beträgt, kann das System kurz nach Eintreten der Situation reagieren.

Die Systemkomponenten wurden sowohl individuell als auch in einem Demonstratorszenario evaluiert. Experimentelle Untersuchungen haben gezeigt, dass die Ergebnisse der automatischen Haltungsbewertung vergleichbar mit denen der manuellen sind.

Des Weiteren war die Methode zur automatischen Haltungsoptimierung in der Lage, die meisten Haltungen mit hoher Belastung in Haltungen mit niedriger Belastung umzuwandeln.

Abgesehen von ihrem Einsatz im WPC können die Grundlagenmethoden, die in dieser Arbeit entwickelt worden sind, auch in anderen Anwendungen im Bereich der Ergonomie und Mensch-Roboter-Kollaboration verwendet werden.

Abstract

The human worker plays a vital role on the production line. Although modern automation technology has succeeded in replacing much of the workforce in factories, there are scenarios where full automation cannot be realised. For example, enterprises offering a broad portfolio of varying products can often not afford the reconfiguration costs of frequently changing manufacturing facilities. Moreover, there are still tasks, such as dealing with highly flexible workpieces, which cannot be effectively automatised. Experienced human workers, in contrast to machinery, are able to adapt to new tasks in a short time and possess cognitive and sensorimotor skills that, at the present time, cannot be imitated by machines. However, maintaining the health of the workforce is a challenge. Work-related musculoskeletal disorders (WMSDs) are the dominant reason for worker absenteeism, leading to large financial losses for states and companies. One fundamental risk factor for these disorders is work that requires awkward postures.

For this reason, extensive research has been conducted into preventing awkward postures in the workplace. The solutions proposed thus far have achieved a significant improvement of the situation. However, implementing the most effectual solutions is labour-intensive and impairs the flexibility of the production line. In brief, robust solutions must be individualised to the specific task and workplace. Although there are approaches that are able to facilitate automatically adapting to new tasks, they are far less effective. In the present thesis, this dilemma is referred to as the "effectiveness-flexibility trade-off".

This thesis strives to bridge this trade-off. Herein, the Working Posture Controller (WPC), a novel type of equipment that imitates the measures performed by workplace designers, was proposed. The WPC automatically monitors the worker's posture when performing a task at hand. In the case of an ergonomically unfavourable situation, the device proposes a re-adjustment of the workplace layout, enabling the worker to perform the same task while maintaining an ergonomically superior posture. The actual re-adjustment can then be realised by an actuator. Here, system concepts and algorithms to produce such a piece of equipment were developed. In detail, the proposed methods can perform posture assessments automatically using a depth camera and optimise posture through workplace layout adjustments. As the computation time needed for these components is just a few seconds, the system is capable of reacting immediately.

The system components were evaluated in isolation as well as in a demonstrator scenario. The experiments revealed that the automated posture assessment results were comparable to a manual assessment. Furthermore, the automated posture optimisation was seemingly able to transform most postures with high physical load into low-load postures.

Apart from their role with the WPC, the fundamental algorithms developed in this thesis can also be applied in other applications in the field of ergonomics and human-robot collaboration.

1 Introduction

1.1 Work-related Musculoskeletal Disorders

The human worker has always played and still plays a vital role on the production line. Despite successes in creating automated systems that are able to outperform humans in carrying out monotonous and high precision tasks, a fully automated manufacturing system is still unthinkable in the majority of cases. In particular, companies with a business model that relies on mass customisation face problems replacing the workforce with technology. The high variance of their products demands a frequent reconfiguration of the equipment, which is expensive and time-consuming. A skilled worker, in contrast, is able to adapt to new tasks in a short period of time. In addition to the problem of limited flexibility, there are tasks which still cannot be automated with today's technology. The human being's sensorimotor skills currently appear irreplaceable.

Although having a skilled workforce is advantageous, maintaining the health of each employee requires additional investment. One of the most serious occupational health threats are work-related musculoskeletal disorders (WMSDs). The European Agency for Safety and Health at Work [14] defines the term musculoskeletal disorders (MSDs) as "health problems of the locomotor apparatus, i.e., muscles, tendons, the skeleton, cartilage, the vascular system, ligaments and nerves". Work-related or occupational musculoskeletal disorders encompass all MSDs that are caused or worsened by work. The most commonly affected body regions are the lower back, the neck, the shoulders, the forearm and the hands. The lower extremities are more rarely involved with a decreased prevalence in the 20% - 30% range [15].

According to statistics, like those from the "Fourth European Working Conditions Survey" [16], WMSDs account for a major proportion of work-related diseases. In a different study, Punnett and Wegman [15] stated that one-third or more of all registered occupational diseases in the United States, Japan and Nordic countries are WMSDs. WMSDs negatively affect both the employee and employer. For the workforce, MSDs results in figures of work absenteeism higher than any other cause. In Germany, WMSDs are responsible for almost one-third of yearly absence days [1] (see fig. 1.1). In the EU, the number of lost days stemming from these disorders is approximately 350 million [7]. As well, for enterprises, WMSDs incur significant costs. According to Bevan [17], health expenses in the EU are estimated at 240€ billion, which translates to up to 2% of the EU

GDP. Falck et al. [18] concluded in their study at an automotive factory plant that poor working conditions led to 3.0 to 3.7 times more quality errors during assembly, which resulted in 6.5 to 7.9 times higher costs to correct them.

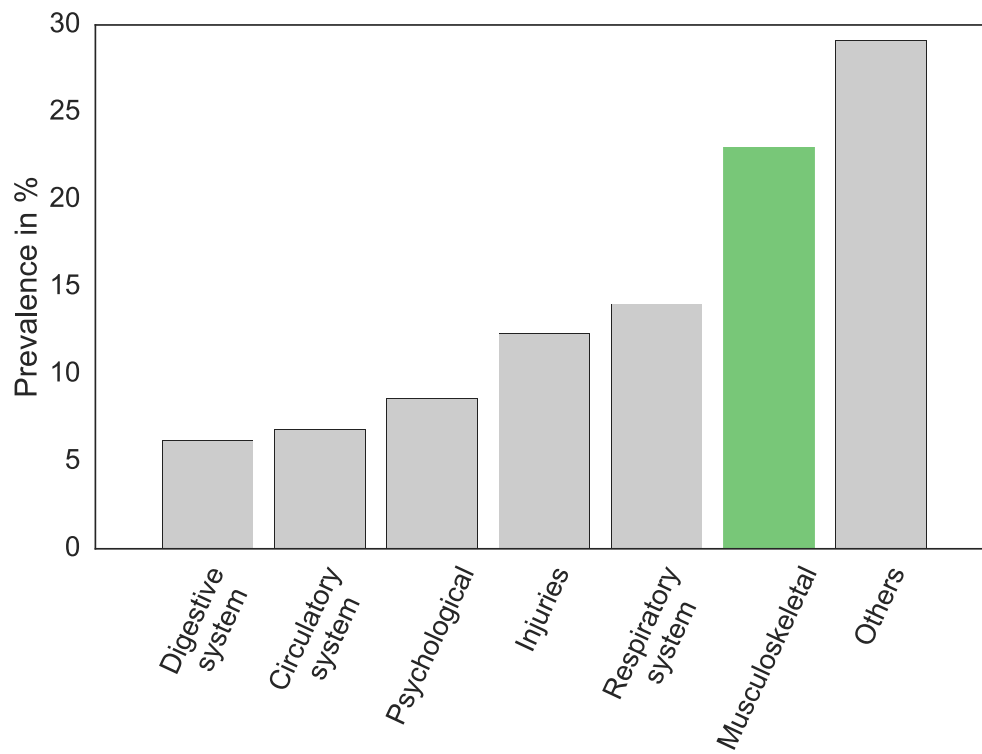


Figure 1.1: Reasons for worker absenteeism in Germany in 2009 according to the German Public Health Fund (AOK) ordered by prevalence (adapted from Badura et al. [1]).

As a consequence of its significant impact, ergonomists, engineers and physicians have searched for solutions to reduce the damage caused by WMSDs. Through their work, significant successes have been achieved. Goggins et al. [5] proposed that on average, the measures implemented have turned out to cover their total costs in less than one year. However, effective measures are work intensive and require expertise. Each has to be tailored specifically to the problem. As an example, Falck et al. [18] put forth that multiple ergonomists, team leaders, quality engineers, specialist operators and production leaders have been involved in their ergonomics improvement processes. Small- and medium-sized enterprises (SMEs) cannot afford such a large number of qualified personnel. Even if the enterprise possesses the required manpower, improving every process becomes expensive and time-consuming. The volume of solutions to be created and evaluated grows exponentially with the number of design options for the solution [19]. To sum up the problem, effective MSD prevention exists and is in great demand, though its price cannot be paid by every company.

1.2 Contribution of this Thesis

Effective WMSD prevention is costly, especially in flexible production scenarios with high process variety. Recent developments in intelligent systems and computer vision have opened up opportunities to outsource parts of the tedious work to computers. Such systems can reduce the required expertise and manpower for WMSD prevention and expedite the measure implementation process.

More precisely, this thesis proposes a system to simplify WMSD prevention. A novel type of equipment was developed that attempts to automatically reduce posture-related WMSDs by usage of intelligent sensors and actuators. It was named the Working Posture Controller (WPC). The WPC continuously monitors the work process and detects situations where the worker has to adopt unfavourable postures to perform the task at hand. Upon occurrence of such an ergonomically unfavourable situation, the WPC adjusts the workplace layout and enables the worker to carry out the same task in an ergonomically superior posture. As well, this thesis developed the algorithmic framework to realise such a device.

The innovation of the present work lies in the realisation and integration of two components - an automated, marker-less in-process posture assessment algorithm and an algorithm to automatically adjust the workplace layout. Through its automated posture assessment and workplace design routines, the WPC attempts to imitate the work of qualified ergonomists. However, its algorithms were designed in such a way that their run-time is only a few seconds in total. Thus, the WPC is able accelerate the measure implementation process, enabling immediate feedback.

To sum up, the scientific contributions of this thesis are:

- The concept for a novel type of manufacturing equipment to prevent posture-related WMSDs was proposed.
- A posture assessment method was put forth that is able to provide immediate feedback.
- A workplace layout optimisation method capable of supplying immediate feedback was submitted.

1.3 Thesis Outline

To present the own work and scientific results, the following structure has been chosen: Initially, a summary of the existing solutions to prevent WMSDs is presented. The result of a literature research demonstrated that there is a scientific gap to be filled - as stated in the introduction, effective solutions to reduce the impact of WMSDs exist, however,

they are costly, work-intensive and require a high degree of expert knowledge. This gap is elaborated on in order to identify specific research goals.

Thereafter, the design of the own solution is outlined. A top-down structure was chosen to explain the details. First, the overall concept of the solution is described. Specifically, there is an explanation how to imagine the device and why it was designed in a particular way. Next, the technical details are covered. Technologically, the solution consists of two main components, and this is why there are two chapters on them. Having explained the whole concept, including the discrete algorithmic details, the prototypical implementation is presented.

To show that the final solution fulfils the defined scientific goals, the evaluation methods and results are presented. Finally, the experimental results obtained are discussed and conclusions are drawn. Overall, this thesis is structured as follows:

- Chapter 2 provides an overview of the approaches proposed by different research areas to prevent WMSDs.
- Chapter 3 summarises the state of the art and reveals the need for a novel type of solution. Moreover, it defines the research questions investigated in this thesis.
- Chapter 4 describes the overall concept of the WPC.
- Chapters 5 and 6 delve into the algorithmic details of each component.
- Chapter 7 covers how a demonstrator system was realised, embodying the WPC concept.
- Chapter 8 describes the process of creating evaluation data and assessments of the system and its algorithms.
- Chapter 9 states to what extent the self-given goal has been achieved. Moreover, it puts forth the strengths and weaknesses of the proposed approach. Finally, the impact this work potentially has on other fields of research is addressed.

2 State of the art

This chapter outlines how different areas of research have proposed to approach WMSDs. It summarises the state of the art methods from different scientific disciplines developed to detect and prevent WMSDs. In general, the prevention requires two problems to be solved. The first involves understanding the problems' source. In this thesis, the causes of WMSDs are referred to as risk factors. In order to implement effective measures, it is essential to understand the underlying exposure types, which are often a composition of different risk factors. This task is also referred as ergonomic assessment. Knowing the risk factors, one has to choose from a variety of measures to eliminate them. The measures are addressed in human factors literature as ergonomic interventions. The second part of the chapter describes the existing manual as well as automated intervention techniques and presents a way to rank them according to their efficacy. Finally, this chapter presents studies that have investigated the effects of various ergonomic interventions and their impact in practice.

2.1 Terminology

Before delving into specific methods, this section introduces terms specifying the words risk factor and exposure. The terms are used throughout the chapter and sketch the focus of this thesis. Winkel and Mathiassen [2] mention three types of WMSD risk factors: physical, psychosocial and individual factors.

The focus here is on physical factors. Most researchers agree that there is a link between physical exposure at work and WMSD risk [15]. Winkel and Mathiassen [2] stated that physical factors exercise their effects through mechanical forces from the body. These forces may then lead to physiological changes. To exactly describe these forces, one would need to express them as a continuous signal containing the magnitude and direction of all force vectors (see Figure 2.1a). In practice, this exposure representation would be too complex. Therefore, the Winkel and Mathiassen [2] proposal was to use a three-dimensional model to describe the exposure signal (see Fig. 2.1b). The dimensions are level, repetitiveness and duration. Exposure level refers to magnitude. Repetitiveness alludes to frequency of level changes. Duration represents the time elapsed during the exposure. Correctly assessing the physical risk factors necessitates documenting all three dimensions.

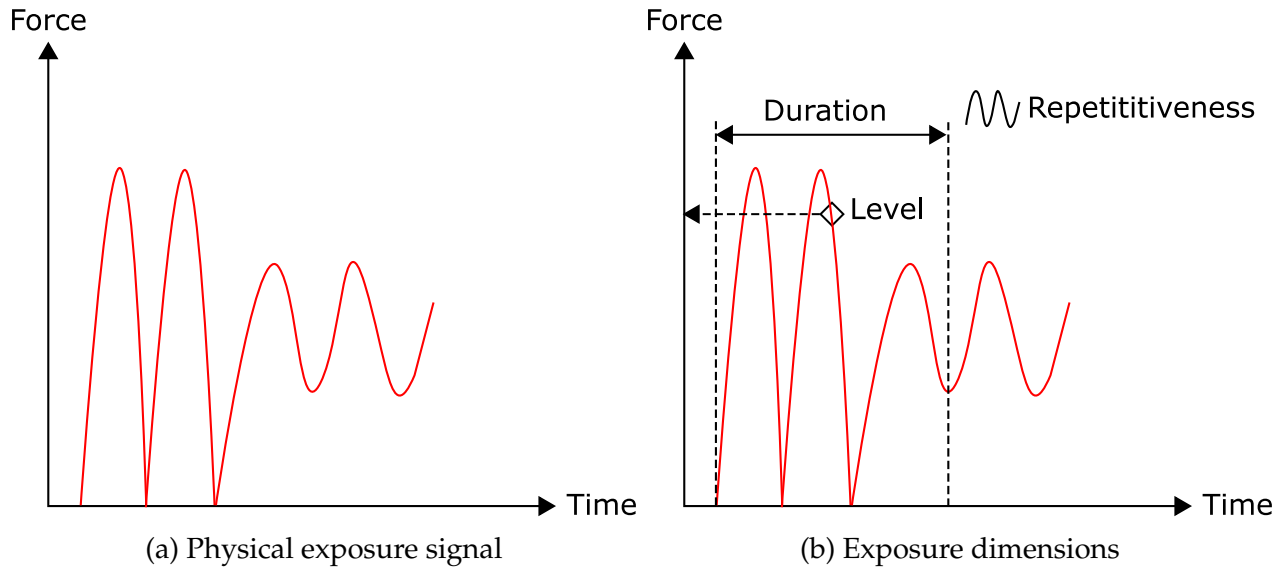


Figure 2.1: Physical exposure dimensions (adapted from Winkel and Mathiassen [2]).

In delving further into physical risk factors, Putz-Anderson et al. [20] mentioned three different cases: force exertion, monotonous strain and awkward working posture. For the sake of simplicity, this thesis refers to these risk factors as force, repetition and posture. The risk factor force is either defined as an external load or as an internal force on a body structure. This factor especially corresponds to the exposure dimension "level". The authors had noted that forceful work imposes a physical load by abduction, flexion, extension or rotation of muscles and other body structures. Repetition characterises monotonous strain originating from cyclical work, which involves repetitive movement. Hence, this risk factor is related to the exposure dimension, "repetitiveness". The WMSD risk is from cyclical flexion, extension, abduction and rotation of muscles and joints. Most studies consider repetition with concurrent risk factors. Therefore, it shall not be considered a primary factor, but always in combination with posture or force [20, 21].

The focus of this thesis is on the risk factor, posture. This factor becomes more prominent when workers adopt awkward static postures for many hours a day, often in combination with forceful work and highly repetitive actions [7]. In practice, we frequently encounter combinations of all three factors. Therefore, this chapter also discusses force and repetition. Figure 2.2 visualises the existing risk factors and exposure dimensions and highlights the focus of this thesis.

2.2 Ergonomic Assessment

The task of ergonomic assessment is to detect potential risk factors at a workplace and assess the exposure dimensions. Thereby, it has to be considered that different workers can

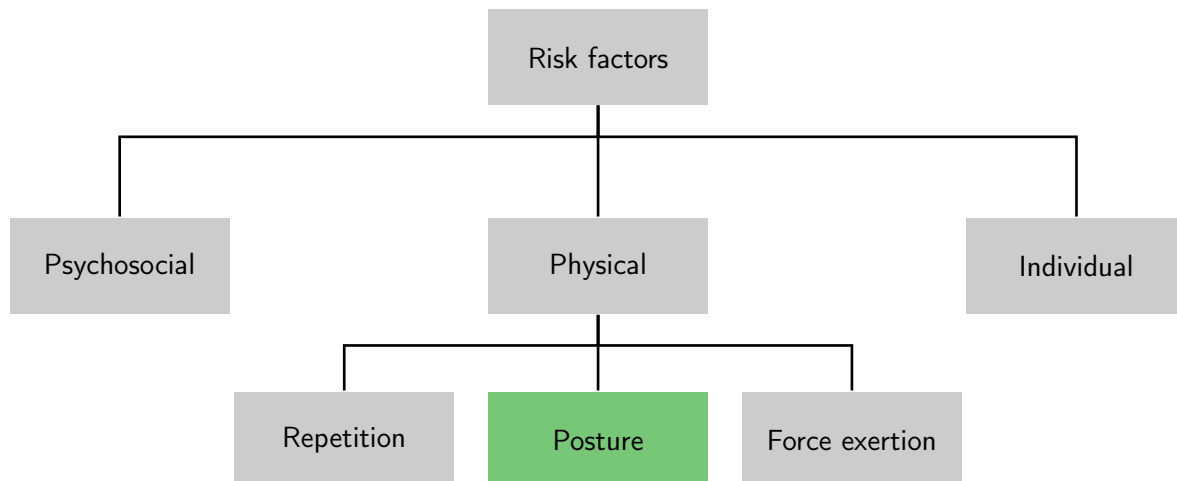


Figure 2.2: Classification of WMSD risk factors and the focus of the current work.

be differently affected by the same exposure. Nevertheless, assessment cannot consider individual physiological responses to the load.

The field of ergonomic assessment is vivid within the discipline of human factors. Researchers have proposed many methods to recognise and assess the WMSD risk factors in the workplace. Li and Buckle [3] and David [22] have both published detailed review papers. The methods put forth can be grouped into three categories: self-reports, direct measurement methods and observational methods (see Figure 2.3). The upcoming sections describe each category in more detail.

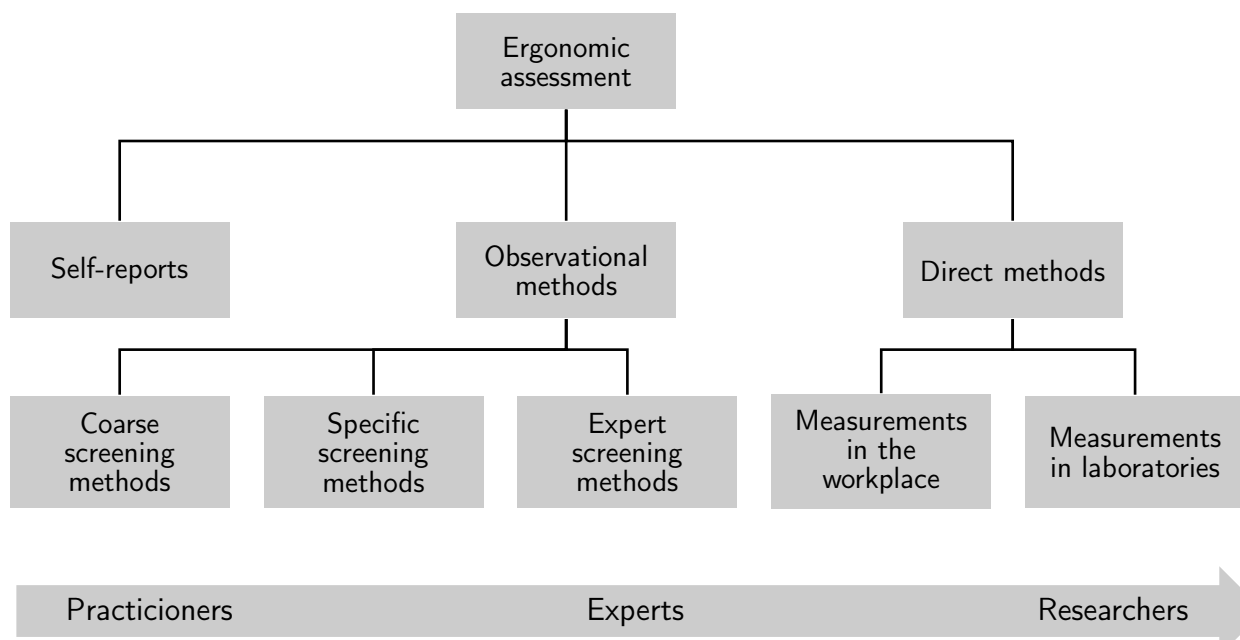


Figure 2.3: Overview of ergonomic assessment methods according to Li and Buckle [3] (second row) and Ellegast [4] (third row). The roles in the arrow denote the user groups the tools are targeting.

2.2.1 Self-Reports

Self-reports assess exposure by asking the worker about their subjective perception of the load while performing the task. This can be done by various methods such as diaries, interviews and standardised and non-standardised questionnaires [22]. Other common methods are rating scales, body maps or check lists [3]. Questionnaires are the most commonly used methods to assess risks due to their simplicity [3]. Moreover, their validity is supported by the worker's approval. David et al. [22] state that the biggest advantages of self-reports are a straightforward usage, the applicability on a wide range of working situations and low cost. In contrast, the result of self-reports can be strongly affected by the individual perception of the situation, lowering the validity and relevance of the statements [3]. Therefore, large sample sizes are needed in order to perform a representative statement. Finally, it is hard to create intermediate results during the task, which prevents realising in-process ergonomic interventions.

2.2.2 Direct Methods

Direct methods express the exposure through physical or geometric quantities acquired from sensors or measurement devices. A summary of the most commonly used tools can be found in Dempsey et al. [23]. They are primarily employed to assess the posture risk factor. The posture can be assessed by Electromyography (EMG) or dynamometers.

Posture assessment methods obtain the data from manual measurement or sensor devices [3]. Goniometers or inclinometers are mechanical tools attached to the body. They indicate the angle between neighbouring parts of the body, and describing a posture can be established by the set of measured angles. Manual devices are inexpensive, but the human operator has to read and note each value by hand. Hence, these devices are only suitable for capturing static postures.

Examples of sensor devices for posture assessment are the Lumbar Motion Monitor (LMM) [24] and CUELA [25] (see Fig. 2.4 for illustration). The LMM makes use of potentiometers whereas CUELA determines the subject's motion by a combination of accelerometers, gyroscopes and potentiometers. The sensors of posture assessment systems are able to yield a continuous stream of data that describes the ergonomic parameters at a granular level of detail. However, most methods require time to calibrate the devices and are susceptible to external interferences. This is impractical for analyses in the factory. Finally, the sensor devices often have to be attached to the body, limiting freedom of movement. This potentially leads to a change in the postural behaviour of the subject [3].

EMG methods measure muscular strain, which can be translated into force. An overview of EMG for ergonomic assessment can be located in Gazzoni [26]. An advantage of EMG is that exposure is directly represented by responses of muscles. However, the measure-

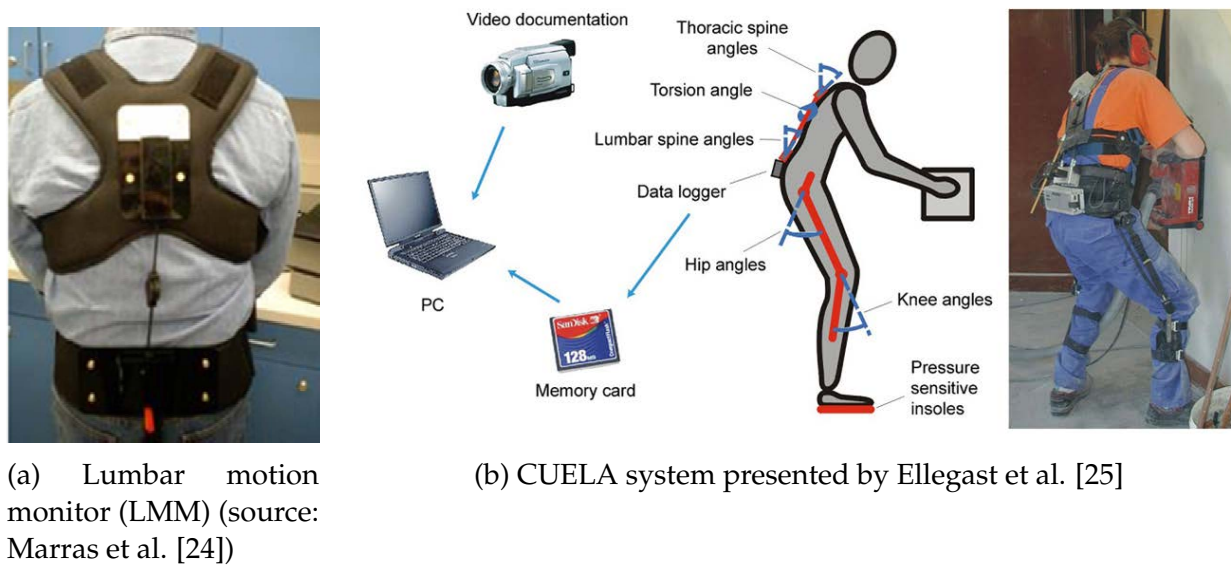


Figure 2.4: Examples of direct methods.

ments are only realisable in a clinical environment based on a tedious calibration process. Furthermore, the signal behaviour is non-linear and varies between individuals, making it difficult to compare the exposure of different muscle groups and workers [3].

Gripping and pinch forces exerted by the hand can be quantified using a dynamometer, e.g., [27]. Dynanometers can measure small forces from light touch ($< 0.05N$) up to power grip forces larger than $500N$.

Briefly, direct methods are a worthy choice for a detailed analysis. However, one must consider the remarkable investment necessary for acquiring and maintaining the equipment. As a vast amount of data is created, the effort required to evaluate the process can be high. Furthermore, based on the tedious calibration routine and the susceptibility to physical interferences, this technology has recently been deemed inappropriate as a monitoring system in a factory.

2.2.3 Observational Methods

Observational methods, as inferred by their name, obtain exposure through observation of tasks. These assessment tools feature a set of pre-defined exposure criteria that a human operator needs to pay attention to. Following the observation of the process, they fill out a worksheet stating whether and how severely each criterion applies. Figure A.2 depicts an exemplary worksheet for postural criteria. Observational methods are primarily used for assessing static posture [3]. However, there are methods that also evaluate force and repetition.

Within the literature, there are various proposed and established observational assessment tools. Examples include RULA [28], REBA [29], OWAS [30], OCRA [31], PATH [32],

LUBA [33] or EAWS [34]. Based on the fact that assessing exposure requires analysing the interplay between multiple risk factors [3], most tools have specific criteria for different factors. Additionally, these criteria distinguish different regions of the body. However, certain methods focus on particular risk factors, such as posture [35, 28, 29, 30], repetition [31] or lifting/material handling [36].

Observational methods can be classified according to their operational principle - they are either event-focused or work sampling-based [37]. Event-focused methods, such as RULA [28], separately assess each event of the process. An event could be the worker adopting a unique static posture or performing a particular action. Often, a process is distinguished by multiple heterogeneous working situations. As such, the time share of a situation influences the overall exposure. Awkward working posture over a greater amount of time increases the risk. These aspects are not taken into account with event-focused methods, which makes them easier to use but produces less nuanced evaluations. To overcome this limitation, work sampling-based techniques [34, 32, 30] have been developed. They consider the process as a set of events sampled over the evaluation period. The final assessment is obtained by accumulating the evaluation results of each event. Although offering more detail, this approach does indeed demand more effort. Table 2.1 classifies the mentioned observational methods according to addressed body regions, risk factors and operational principle.

Table 2.1: Summary of observational methods.

Method	Body region	Risk factor	Operation principle
RULA [28]	upper body	posture, force	event-focused
REBA [29]	whole body	posture, force	event-focused
OCRA [31]	upper limbs	posture, force, repetition	event-focused
LUBA [33]	upper body	posture	event-focused
NIOSH [36]	not specific	force (lifting)	event-focused
PATH [32]	trunk, arms, lower body, neck	posture	work sampling-based
OWAS [30]	trunk, arms, lower body, neck	posture	work sampling-based
EAWS [34]	whole body	posture, force, repetitionx	work sampling-based

Ellegast [4] differentiated observational methods by their level of complexity. The methods were placed into four unique categories: "Grob-Screening-Verfahren" (coarse methods), "Spezielle-Screening-Verfahren" (specific methods), "Expert-Screening" (expert methods) and "Messverfahren" (direct methods). The methods are recommended based on the different user groups. Ellegast distinguished the user groups as practitioners, experts and researchers. Figure 2.3 (bottom row and arrow) portrays a visualisation of Ellegast's taxonomy. The intention was to reduce the assessment effort by first ap-

plying the inexpensive but rough methods. Afterwards, the finer but more expensive methods are employed to further analyse the "hot spots". In the following, each level of complexity is roughly characterised.

Coarse methods are used to quickly obtain an ergonomic overview. They are the simplest and most time-efficient way to evaluate the exposure. Unfortunately, they do not provide a numeric risk score, rendering them too inaccurate to monitor the effects of workplace design [3]. When risk has been detected by coarse methods, specific methods or expert methods can be utilised for a more sophisticated assessment of these problems. Specific methods offer fewer details, though they are easier to use in comparison to expert methods. Expert methods offer a higher level of detail at the cost of more complexity for the user. Ellegast [4] noted that a clear distinction between both groups is hard. Rather, there is a smooth transition between specific and expert methods. An advantage of both is that they express risk as a numeric value that represents either the specific physical load for different risk factors or the overall load. Such a score is necessary to quantify the effect of ergonomic interventions [3].

Although many factors can be quantified and detected by expert methods, there are problem cases that cannot be satisfactorily covered. These are namely 3D postures or asymmetry effects involving torsion or lateral bending (see Figure A.4) and motion over time. This is where direct methods come into play - they offer the greatest level of detail with a high cost investment. Additionally, as described previously in Section 2.2.2, direct methods often have a limited application area and are principally used in the laboratory.

In general, the biggest advantages of observational methods are that they are inexpensive to apply and can be used in industrial as well as clinical environments. Yet, their reliability and their ability to precisely predict risk remains unclear [3]. The weighting of each risk factor when combined into an overall result is often described as "hypothetical". Multiple trials with different subjects and analysts are needed to overcome the subjective bias of such results. Seeing that the subjective estimation variance is especially high in dynamic movements, these methods are mostly employed for assessing static tasks. Finally, despite the many efforts to optimise workflows, observational methods continue to be work intensive because they require a manual analysis of the process or the video material.

2.2.4 Summary of Ergonomic Assessment Methods

An ergonomist can choose from three ergonomic assessment method families: self-reports, observational methods and direct methods. These tools differ with regards to cost, versatility, generality and exactness [2].

Self-reports are a cost-efficient means to assess various risk factors - only questionnaires are used. As the design of questionnaires is not standardised, one can create individualised assessments concentrating on the features of interest. The disadvantage, however,

is that self-reports are highly subjective - the data collection is performed by the subjects themselves, leading to a limited validity of the results.

Direct methods are a stark contrast to self-reports. The equipment needed to conduct the studies is quite expensive. Direct methods focus on specific details of the risk factors that can be measured through physical quantities. The consequence is that risk factors, which cannot be exactly represented by measurable physical quantities, cannot be assessed, significantly restricting generality. However, the results produced are exact and objective if the relationship between risk factor and physical quantity is valid.

Finally, observational methods represent a compromise between the aforementioned extremes. The cost is moderate - expensive specialised equipment is not required. However, trained specialists are needed to perform the data acquisition task. There are a variety of different observational tools (see Tab. 2.1). The aspects assessed in these tools are analysed in a standardised way, which does not allow an assessor to incorporate additional features. Hence, the versatility of these tools is lower than in self-reports. As well, because the assessment for different subjects is chiefly performed by one or a few specialists, the results are expected to be more exact than those from self-reports. Nevertheless, it is worth noting that the assessments of different specialists are subjective. Figure 2.5 summarises the advantages and disadvantages of various ergonomic assessment methods.

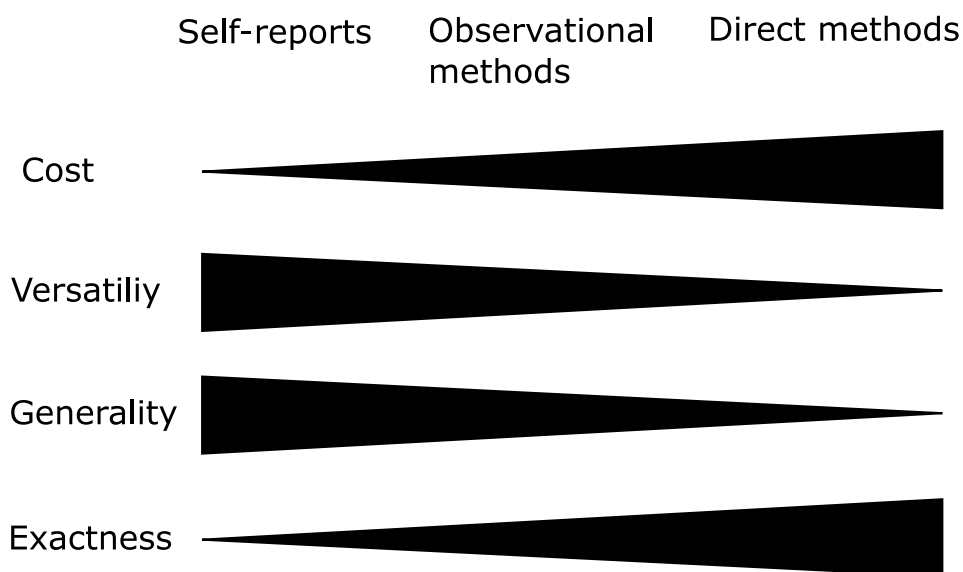


Figure 2.5: Summary of ergonomic assessment methods (adapted from Winkel and Mathiassen [2]).

2.3 Ergonomic Intervention

Having gained insights into which risk factors are present and their extent, appropriate measures can be initiated. Ergonomic intervention is a broad field of research where countless studies have been published. Describing each method in detail would be outside the scope of this work. Rather than providing an extensive review, the purpose of this section is to present a taxonomy classifying intervention techniques from human factors and engineering according to their effectiveness and mention examples. The main goal is to derive a "dead spot" where further research is necessary.

Review papers on ergonomic intervention techniques can be found in [38, 39, 40, 21, 41, 42, 43]. The methods can be classified according to the risk factor they address. Moreover, Van der Molen et al. [38] distinguished the categories as: technical, organisational and individual measures. Technical measures refer to the design of workplaces or the introduction of equipment and aids. Organisational measures involve job scheduling and actions changing the composition of a worker's tasks. Individual measures mean training and education of the workforce. Similarly, Bergamasco et al. [40] used the categories, structural measures (e.g., workplace design), organisational measures and training.

Considering the variety of approaches, Goggins et al. [5] proposed a categorical scale to rank the effectiveness of intervention methods. This guide is intended to simplify the choice of a method for planners and engineers. The scale comprises four levels - methods that eliminate the exposure (level 1), methods that reduce the level of exposure (level 2), methods that reduce the time of exposure (level 3) and methods where success relies merely on the behaviour of the worker (level 4).

At the outset, methods are preferred that eliminate exposure (level 1). Examples are automating or mechanising the exposure-intensive parts of the task or designing them out through workplace layout arrangement. If an elimination is not feasible, interventions are opted for, which reduce the level of exposure (level 2). One example is diminishing weight to be lifted. At the next-best level, there are methods which reduce the duration of exposure (level 3). This can be attained by job rotation where tasks are arranged such that they contain different amounts of exposure or no exposure at all. This avoids monotonous strain of particular body structures. Finally, if all the aforementioned measures are not feasible, one has to rely on methods that require the active participation of the worker (level 4), such as worker training. The downside of this intervention type is that its effect is dependent on the worker's behaviour and their willingness to constantly apply the measures. Figure 2.6 summarises the categories.

The following sections (2.3.1 - 2.3.3) present exemplary conventional intervention methods. They primarily rely on human actions and are classified according to the risk factor they address. Subsequently, there is a closer look at methods using computers and automation technology in Sec. 2.3.4, as this topic is a core element of the present work.

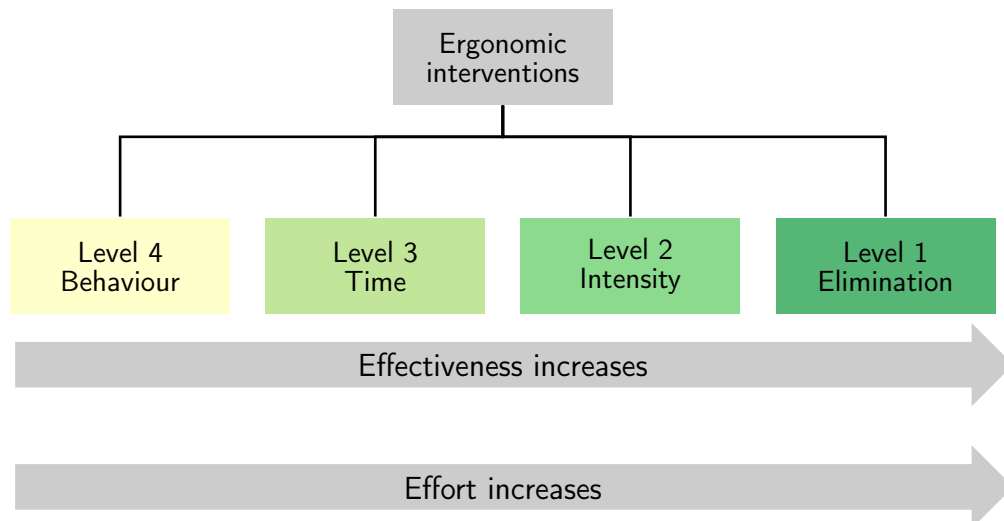


Figure 2.6: Classification of ergonomic interventions according to Goggins et al. [5].

As a number of computer-aided techniques cover multiple risk factors or only support conventional measures, they are not ordered by risk factor.

2.3.1 Interventions against Force Exposure

To reduce the severity of the risk factor force, the components of a job where force exertion exceeds an acceptable level have to be separated from the worker. Alternatively, the exposure level can be reduced to an acceptable value. Schaub et al. [44] empirically determined the maximum recommended force limits for each force direction and thus provided a clear definition for what is meant by "acceptable".

The complete separation of the worker from the force (level 1) is accomplished by structural measures, including automation or mechanisation of the forceful tasks. Schmidtler et al. [45] called the devices used lifting aids or transport assistance. Mechanical lifting aids have been previously introduced in earlier studies, such as Marras et al. [46].

Methods that alter the process design are considered level 1 - 2 methods. For example, decreasing the weight of objects to be handled [38] can either abolish the risk factor or at least diminish the exposure level. Exposure time (level 3) can be reduced by allowing longer breaks or by introducing job rotation. Finally, ergonomic interventions have often featured lifting technique training [47] to attenuate strain on critical parts of the body (level 4).

2.3.2 Interventions against Repetition

For the repetition risk factor, a full elimination (level 1) can only be reached by significantly lowering the number of repetitions or completely automating the process. Bergamasco et al. [40] put forth level 1 - 2 methods from the group of structural measures. These

are namely workplace layout optimisation and the introduction of aids. The purpose is to decrease the utility of force, awkward postures and general movements. Furthermore, there are organisational measures (level 3) to lower task frequency or time or to increase rest periods. Job rotation is a prominent member of this group. Moreover, worker training (level 4) can be employed to address the correct motion while performing the task. Additionally, equipment can be introduced that encourages the worker to take a break after a pre-defined duration of working time [48]. Although automation is involved in this measure, it is considered as a level 4 method, as the human being has to ultimately perform the action.

2.3.3 Interventions against Awkward Posture

The key to eliminating the postural load is to enable the worker to adopt a "neutral" posture. With a neutral posture, the joints are not bent and the spine is aligned and not twisted. It places the least physical load on the body. The method of choice is workplace layout design. In most cases, workplace geometry has to guarantee that ergonomic guidelines apply for a range of anthropometric properties [49]. If a single design cannot cover all anthropometric variants, lift tables and other geometric adapting devices can be employed [50]. In the best case, the layout design can completely do away with the risk (level 1). Otherwise, it can at least diminish exposure intensity (level 2).

Another interesting level 1 - 2 approach comes from Weidner et al. [6], who established the concept of human-hybrid robot (HHR). Their main idea was to attach technical kinematic chains to the body to transmit the force from vulnerable parts of the body to more robust areas. The kinematic chains have a modular structure such that they can be combined and tailored depending on the task to be supported. In their article, an exemplary use case is described where a module supports the arms for overhead work.

By employing job rotation schemes, the time spent in awkward postures can be decreased (level 3). Finally, training the subject to regularly alternate between postures (level 4) can aid in reducing risk. As a consummate guideline, Neuhaus et al. [51] encouraged the worker to sit half of the time and stand the other half.

2.3.4 Computer-Aided Ergonomic Intervention

Sections 2.3.1, 2.3.2 and 2.3.3 have mostly presented intervention measures, all of which are manually executed. This section covers interventions that use computer technology as they are relevant to the present work. Ergonomic interventions purely reliant on human beings have drawbacks. First of all, ergonomic assessment and intervention design are time-consuming tasks. Assessing ergonomics with self-reports and observational methods require tedious manual analysis of the acquired data. In addition, properly designing workplaces is time-consuming because it involves iteratively testing multiple prototypes

[7]. Workplace designers are frequently experts with specialised training. Even when the solutions developed are effective, they are indeed costly. Furthermore, level 4 intervention methods primarily depend on the behaviour of the worker, and if they do not consistently apply the techniques learned, the desired effect is weakened.

With the development of computers and software algorithms, systems can be realised that are able to support ergonomic interventions. Through heavy processing power, they can even perform human tasks in significantly shorter amounts of time. Herein, three examples of how computers, sensors and actuators have been integrated into ergonomic interventions are discussed. The technologies are digital human models, in-process feedback systems and human-robot collaboration. These examples have been chosen as they are the most similar to the present work.

Digital Human Models and Simulation Tools

DHMs are a virtual representation of the real worker integrated into factory simulation tools. DHMs allow the designer to simulate the human in a virtual workplace to perform task-related as well as ergonomics-related analyses. Using DHMs, the risk factors of force, repetitiveness as well as posture can be identified.

Delleman et al. [7] gave four reasons for why these techniques were created. First of all, DHMs offers an opportunity to reduce development costs. Designing and producing hardware prototypes is laborious and cost-intensive. When multiple physical prototypes have to be evaluated, this means the traditional design process becomes even more expensive. Moreover, these costs rise when designs are incrementally improved. While requiring greater effort to digitise the prototypes, DHMs pays off during the later stages of development (see Fig. 2.7). Secondly, verifying prototypical workplaces necessitates a large population with different anthropometrics and capabilities. DHMs provide an efficient way to test the setup with different virtual individuals. Thirdly, the 3D representation of the DHM enables additional visualisations. For example, a worker's field-of-view can be displayed, which is hard to obtain using traditional assessment methods. Finally, DHMs feature a common platform for various specialists, e.g., physicists, engineers or experts in human psychology, to contribute to workplace design. Hence, it facilitates a common means of communication.

Meanwhile, DHMs have matured to such a state that they are widely used in the planning of manufacturing processes. With this, software developers have released a variety of commercially available DHM tools. Examples include AnyBody [52, 53, 54], Santos Human [55], RAMSIS [56], Jack [57] and ema [58]. The DHM tools are employed in the design, modification, visualisation and analysis of virtual workplaces or product interactions [59]. Typical tasks are determining maximum reach, clearance or analyses of the line-of-sight and work ergonomics [7]. DHMs can be utilised for analyses in prospective (designing a new workplace from scratch before the physical prototype exists) or in retrospective (ergonomic analyses of existing workplaces) scenarios. The ability to simulate multiple "what if" scenarios with different worker populations in an acceptable amount

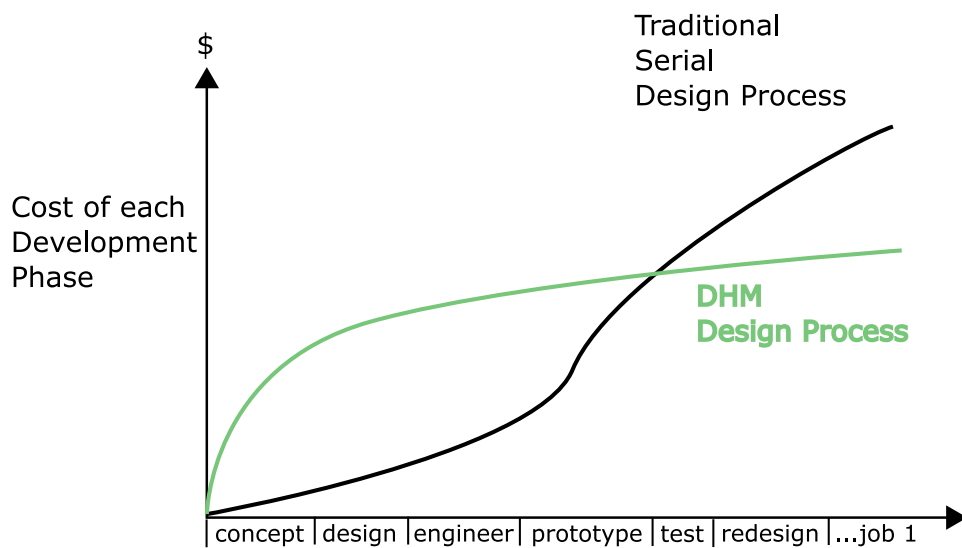


Figure 2.7: Comparison of traditional workplace development costs with DHM workplace development costs (adapted from Delleman et al. [7]).

of time [7] significantly reduces the costs of the planning process. Exemplary ergonomic simulations and assessments have been conducted for PC assembly [60], manufacturing work cells [61], car assembly processes [62, 63], railway coach assembly workplaces [64] and human-robot collaboration workplaces [65, 10]. Additionally, ergonomic assessment and optimisation of products can be accomplished through DHMs. As a case in point, Rasmussen et al. [52] have analysed and optimised the forces coming from the usage of a saw with their tool, "AnyBody®".

Besides ergonomic assessment, DHMs have been featured with methods to conduct automated workplace layout design. Abdel-Malek et al. [8] expressed the layout design task as a mathematical optimisation problem where target points are specified in the working environment. For given workplace geometry and motion ranges of the DHM, a number of objects, like boxes, workpieces or tool holders, have to be located such that certain ergonomic criteria are fulfilled. The solution to this problem can then be determined through genetic algorithms [66], non-linear programming [67] or Monte Carlo simulations [68]. Lämkuhl et al. [59] stated that DHMs are increasingly becoming the major means for the ergonomic analysis of assembly processes. However, there are certain challenges that still need to be tackled in order to make this tool more effective. Firstly, using DHMs requires expertise. Engineers working with DHMs have to possess skills in ergonomics as well as CAD modelling. One significant reason for the grand development costs in the concept, design and engineering phase (see fig. 2.7) is the amount of time required to model a realistic behaviour of the DHM during the process. Accurate programming of the DHM motion and behaviour is tedious. A DHM contains a two-place number of joints, which in themselves contain three coordinates. Programming a motion corresponds to setting the position coordinates for each joint at each point in time.

Thereby, kinematic constraints of the body have to be considered and it needs to be guaranteed that the DHM does not collide with objects in the digital environment.

There have been numerous attempts to reduce the complexity of DHM programming by motion capture and automated motion generation methods. The motion capture system observes a human subject performing the working motion and tracks the positions of the limbs. Normally, the test subject has to wear markers that are then recognised by the motion capture system. Thus, the programming of the DHM behaviour is by simply demonstrating the motion. It is either applied directly in the process or via a mock-up workplace. Härtel et al. [69] applied a Vicon® motion capture system to transfer motions to the DHM Alaska/Dynamicus [70]. Similarly, Daphalapurkar et al. [71] developed a multi Kinect® system that interfaces with the Jack DHM [57]. Unlike expensive marker-based systems, the Kinect® sensor operates without markers but with a compromise of reduced accuracy and robustness.

The idea behind the motion generation approach is to manually define high-level tasks and let the computer generate specific low-level motions. Thus, the behaviour can be programmed without caring about the motion in detail. There are two options to create the motion - either fundamental motions are pre-programmed and assembled into a high-level task or they are generated by Inverse Kinematics. When pre-programming fundamental motions, the main effort is invested in creating realistic small motion blocks, which can then be assembled to create process motions. Fundamental motions are then often recorded by accurate motion capture techniques supported by biomechanical studies to validate the motion for different environmental conditions. However, the creation of a fundamental motion requires a large number of samples in order to be able to represent the variance between individuals. As an example, Manns et al. [72] have estimated a lower bound of 149 motion capture takes for a simple picking motion.

A way to escape this data requirement is to incorporate assumptions and simplifications regarding the motion that concern the environment and relevant biomechanics. Mårdberg et al. [73] expressed the space of feasible motions as a formal language. Each high-level task is defined as a composition of low-level motions. In turn, each motion needs a particular environmental state and manipulates the environment in a unique way. A human operator only has to define the sequence of tasks. Employing the task definitions, the algorithm automatically searches for a feasible sequence of motions that always satisfies the environmental conditions. Furthermore, a manually generated sequence can be formally verified to guarantee that the motions do not contradict one another. For instance, an object can only be grasped if it lies within the target place. To demonstrate their idea, the authors applied the formal language motion generation approach on a cable assembly use case. Nishiwaki et al. [74] proposed an approach for motion generation specific to humanoid robots. With their method, reach motions for robots were produced. The main idea was to have pre-programmed basic reach postures that were adapted to an unknown scenario. The generation of the new motion involved selecting a suitable pre-programmed posture and modifying it using inverse kinematics. Additional steps, such

as balance compensation, were introduced to make certain the robot had a firm stance. Kuo and Wang [75] utilised the Methods Time Measurement (MTM) system as a basis for generating the motion. The MTM defines a set of elementary motions which can be combined to describe working tasks. A typical use case of MTM is to measure the required cycle time for a task by summing up the expected times for the elementary motions. The authors first converted a task represented by MTM-UAS codes into corresponding motion templates. The templates were constrained by inputting the task type. Afterwards, the motions were refined using Inverse Kinematics. Finally, a user could manually edit the generated motions and correct small errors therein. The DHM ema [62] creates motions from small building blocks representing common operations. Examples are sitting down, walking or picking up an object [76]. The fabrication of realistic building blocks is supported by biomechanical studies and motion capture devices. The blocks can be parametrised to individualise the motion to a particular task, such as an object pick that can then be manifested using the right or left hand.

The Inverse Kinematics approach is used for more low-level motions where joint trajectories have to be adapted. The underlying rationale is to diminish the programming effort by only defining target points or trajectories for particular joints. Additionally, constraints for the trajectories, like minimal and maximal joint angles, can be added into the model. The other joint trajectories are then computed through solving an optimisation problem. The realism of the motion depends on how well the constraints on the motions represent the reality. Miyata et al. [77] used Inverse Kinematics to generate lift motions. First, they recorded the trajectory of the hand using motion capture techniques. Afterwards, the whole body motion was generated by solving a control problem of a redundant manipulator with a given end-effector trajectory. Similarly, Hue et al. [78], Fourquet et al. [79] and Abdel-Malek et al. [80] modeled the motion generation problem as an optimisation problem. The algorithm takes target points for limbs as the input and outputs a motion or posture that moves the target limbs to the target points. The objective function considers the distance between limbs and target points as well as additional constraints on the joints and trajectories.

Despite the motion generation techniques integrated in DHMs, the realism of the simulation results remain questionable. Lämkuhl et al. [59] performed a study for 155 use cases where ergonomic assessments with DHMs were compared with manual ergonomic analyses. Their findings were that in about 48% of all cases, both approaches were in agreement. These were mostly processes containing standing or unconstrained working postures. Unconstrained postures comprised tasks without external forces or weights. Thus, working height optimisation tasks could be supported. Nevertheless, the simulation accuracy for constrained postures or dynamic motions still require further improvements. For example, balance prediction has been considered a problematic use case.

In-process Feedback Systems

In-process feedback systems, or alert systems, are referred to as tools which acquire data from ergonomic assessment tasks and are able to supply immediate feedback to the user. In-process feedback systems, for the most part, determine and assess working postures. Either image-based methods [81] or acceleration/inertia sensor-based technology [82, 83, 9] have been employed for ergonomic assessment. While image-based methods do not restrict freedom of movement, they have only been tested in controlled environments. In utilising Inertial Measurement Units (IMUs), in contrast, the posture assessment is robust enough to be applied in an industrial environment. However, they require wearing bulky equipment. The data acquired can be used for further analysis by specialists [83].

Martin et al. [81] presented a system to teach workers the effect of their lifting and carrying techniques. It uses the Microsoft Kinect® sensor in combination with its joint localisation algorithm [11] from the Kinect Software Development Kit. The postures are evaluated according to the OSHA model [84], and the system has been evaluated on various exemplary lifting postures in a laboratory environment.

Vignais et al. [9] created a system that performs in-process ergonomics assessment in industrial environments (see Figure 2.8). It utilises IMUs to submit an analysis according to RULA [28]. Visual feedback is provided by a head-mounted display depicting which limbs are in a critical posture. The authors conducted a user study that came to the conclusion that the system was able to reduce the risk score while simultaneously not significantly increasing execution time of the studied process.

In Lins et al. [83], wearable IMUs were employed to assess the posture of the upper body. The emphasis of their work was on designing sensor networks that limit the freedom of movement as little as possible. The authors saw the most important application during the early detection and prevention of gonarthrosis and lower back pain. The data acquired by the wearable was analysed by physicists or used for warning purposes, with which the user could then attempt to actively alter the posture or, if that was not possible, take a break.

Human-Robot Collaboration

Human-robot collaboration can reduce WMSD risk by delegating the forceful and monotonic parts of a task to robotic manipulators, so-called intelligent assist systems (IASs) or intelligent automation devices (IADs) [85]. With elaborate safety mechanisms, human beings and robots can share the workplace, combining human flexibility and dexterity with robots' force and accuracy. In this work, two types are distinguished:

The first type uses a separate robot manipulator [86, 85, 87, 45, 10]. A sub-category of this group are Collaborative Robots (Cobots), introduced by Peshkin and Colgate [86]. Cobots are handling systems that combine the characteristics of industrial robots and handheld manipulators [45] (see Figure 2.9a for an example). They are haptically guided by a human force exerted onto a control interface. Their task is to provide power support

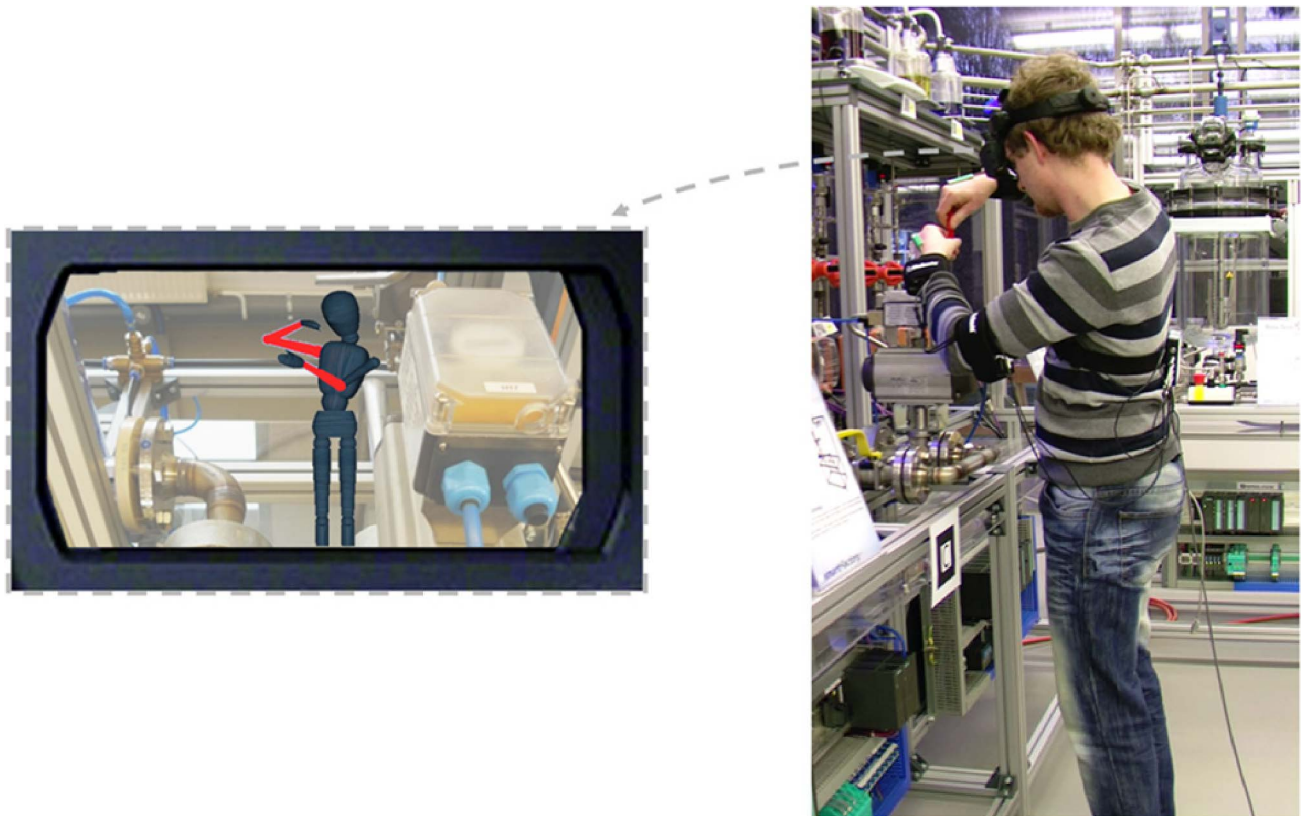


Figure 2.8: Visual feedback for working postures proposed by Vignais et al. [9].

and guidance [85, 45] to the worker in handling tasks, like lifting or transporting heavy objects. In addition to force reduction, cobots possess advanced safety mechanisms for avoiding collisions with the environment.

A particularly interesting approach was presented by Busch [89]. An assistance system consisting of multiple robots was produced to decrease the labour-intensive manual handling of heavy parts during welding. The idea was to employ the robots in holding the parts in such a manner that the welder could process it without the need to adopt awkward postures. Based on the shape of the workpiece and the welding task, robot motion could be programmed offline which permitted the worker to adopt a neutral working posture. Motion planning was possible with the robot cell planning software, FAMOS, of Carat Robotic Innovation GmbH [90]. The character animation software, EmotionFX, by Mystic Game Development [91] was employed to model realistic human motion during the task. Although this simulation approach provides a lower level of detail in comparison to established DHMs, it represents a cost-effective and sufficiently accurate solution to plan the workplace. To assess ergonomics in each scenario, the tool, OWAS [30], was used. The complete exercise was carried out when planning the workplace as a pre-process intervention (see Fig. 2.10 for illustration). In contrast to Cobots and other handling aids, this system specifically targets the posture risk factor.

The second group of human-robot collaboration solutions are exo-skeletons, designed as external mechanical structures (see Figure 2.9b for an example). Its joints correspond to

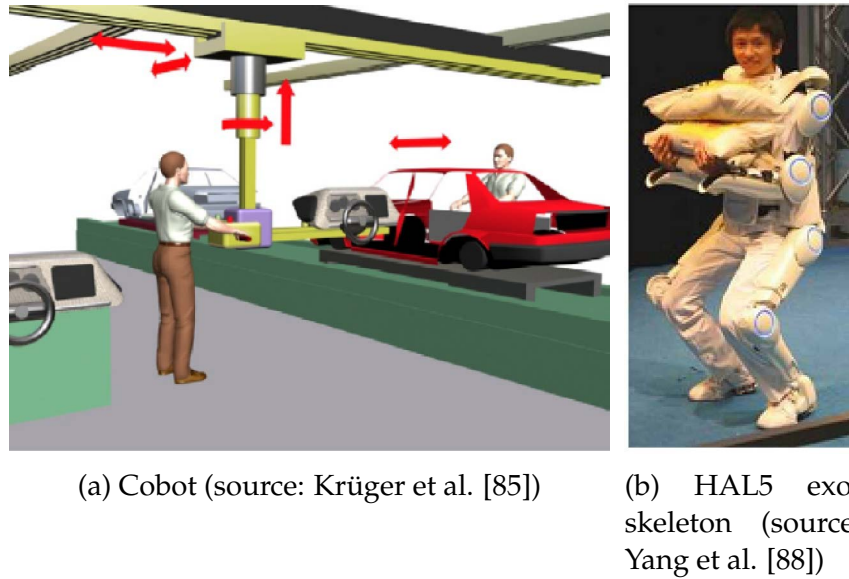


Figure 2.9: Illustration of Cobots and exo-skeletons.

those of the human body [88]. Weidner et al. [6] presented a modular exo-skeleton system for assembly tasks. Their work began with the introduction of kinematic chains that provided force support and reduced the postural exposure through transmission of the force to less vulnerable parts. Active modules, which operate similar to traditional exo-skeletons, were the core component. Additionally, there were passive modules, mostly for force transmission, which acted as more cost-efficient alternative. The main advantage of this technology is that the modules can be arbitrarily combined and tailored to the process, making the solution flexible for different use cases.

Similar to Cobots, the main duty of exo-skeletons is to provide power support for the worker while maintaining safety. The major benefits of them is that they are designed to be mobile, thus being able to supply support where needed. However, this comes with the challenge of an autonomous energy supply and greater complexity in controller design along with adaptation to the human body [85, 45]. Apart from industrial workplaces, their potential has already been applied in medical areas, such as rehabilitation [92].

In order to correctly provide support, the worker's intention has to be accurately understood within a short period of time. This is mostly possible with force-torque sensors. Force profile data can be employed for haptic feedback, making the task supervision component for humans more intuitive.

2.3.5 Impact of Ergonomic Interventions

As described throughout the chapter, researchers have proposed variety of ergonomic intervention measures. There have been a number of approaches to estimate the effectiveness of these methods. Nevertheless, a theoretical scale on its own does not necessarily

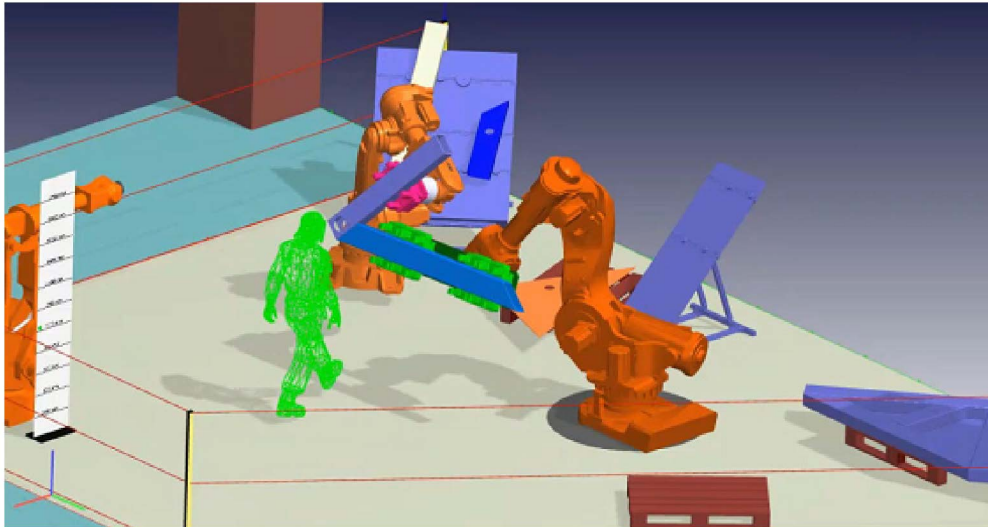


Figure 2.10: Safety and ergonomics simulation of a robot cell described in Thomas et al. [10].

predict impact in practice. Hence, in the literature, there are studies that have reported various ergonomic interventions in practice. The objective of this section is to summarise these results and answer the following questions:

- What are effects of ergonomic interventions? Do they pay off?
- Which types of interventions achieve the greatest effect in practice?
- How can a managing board of a company be persuaded to invest in ergonomic interventions?

Goggins et al. [5] proposed a scale for measuring the effectiveness of ergonomic interventions (see Sec. 2.3 for a detailed explanation). This scale was derived from a predictive cost-benefit analysis, of which the results were published in the same article. The analysis included 250 studies from ergonomics programmes in offices as well as in industrial environments. The authors conducted an evaluation before and after the intervention and considered various performance measures, such as WMSD incidence rate, absence days or compensation costs. In almost all cases, dramatic improvements were documented. For example, incidence rates dropped by 65% on average and companies had to pay 68% less the original compensation costs subsequent to the intervention. Additionally, productivity of the workforce and the quality of the products was enhanced. The remarkable fact was that the benefits were realised in less than a year.

Van der Molen et al. [38] concentrated on analysing the long-term effectiveness of manual handling tasks. This systematic review included 44 studies, and, like Goggins et al., improvements in working conditions and WMSD symptoms were noticed quite quickly. The most significant reductions of physical load were achieved when lifting devices were employed. Furthermore, most studies described an improvement in worker behaviour - participants started to actively circumvent exposure from bad techniques.

Keyserling et al. [50] studied an automotive corporation, evaluating 151 jobs in total. The goal was to develop an intervention program for avoiding awkward postures. Every six months, an evaluation was performed. Overall, the results indicated a successful diminution of awkward trunk and shoulder postures. However, interventions for neck postures were less effectual. Out of all intervention types, workplace design demonstrated the highest efficacy.

Engels et al. [47] evaluated ergonomics education courses in nursing. Awkward postures at work were assessed for 12 subjects before and after the intervention. Additionally, there was a control group that comprised 12 subjects that did not receive an educational course. The authors observed that the group which received educational courses had a significant decrease in adopting a number of harmful postures and committing errors versus the control group.

Falck et al. [18] looked at the impact of ergonomics on quality in car manufacturing. The assembly quality of over 24,000 cars was examined. Their results showed that there was a strong correlation between ergonomic load and quality, as well as cost. Medium and high physical load assemblies increased the risk of an assembly error by 3- and 3.7-fold, leading to 8.2- and 8.7-fold more total action costs than low physical load tasks.

Despite many investigations verifying the positive effects of ergonomic interventions, there are also studies that did not describe a significant improvement. Driessen et al. [48] performed a review on ergonomic interventions for averting lower back and neck pain. The authors chose only studies containing randomised controlled trials. In total, 10 studies were analysed, and there was not much evidence that an intervention significantly decreased neck pain intensity compared to cases of no ergonomic intervention.

In brief, there are various studies that observed positive effects of ergonomic interventions as well as those that have not demonstrated their merits. However, over the course of the literature research for this thesis, papers from the first group dominated by a substantial margin. In addition, it has to be noted that the papers where it was not possible to definitively establish improvement took into account fewer studies than those featuring positive reviews.

2.3.6 Summary of Ergonomic Intervention Methods

Ergonomic interventions constitute a vibrant field of research. Studies have shown their positive impacts on worker health, productivity and costs. However, not all measures are equally effective. The aforementioned ergonomic intervention techniques are summarised in Tab. 2.2 and classified according to the risk factor they address and the best level of effectiveness they can accomplish. Note that depending on the specific implementation, certain methods can belong to different levels of effectiveness.

Table 2.2: Interventions classified according to effectiveness and addressed risk factor.

Level	Posture	Force	Repetition
1 - eliminate	Workplace layout design, HHR modules	Process design e.g. reduce weight, Cobots, Exo-Skeletons, Lifting aids	Process design e.g. automatisisation
2 - intensity		Process design e.g. reduce weight	reduce force and postural load per cycle
3 - time	Job rotation	Job rotation	Job rotation
4 - behaviour	In-process feedback systems, training e.g. alternate posture	training e.g. lifting technique	In-process feedback systems, training e.g. technique

An interesting fact of ergonomic intervention techniques is that they either require preparation before the worker performs the task (e.g., when planning the workstation) or can be directly applied during the task - either by the worker or by equipment. Examples for the first category include workplace design techniques or worker training. In the following, these methods are referred to as pre-process techniques. Cobots, lifting aids or alert systems are examples for the second category, referred to as in-process methods. Table 2.3 lists examples of both pre-process and in-process measures. In brief, pre-process in-

Table 2.3: Interventions assigned to either pre-process or in-process measures.

	Pre-process	In-process
Posture	Workplace layout design, Job rotation, Training, HHR kinematic modules	In-process feedback systems
Force	Process design, Lifting technique training	Cobots, Exo-Skeletons, Transport assistance
Repetition	Process design, Job rotation	In-process feedback systems

terventions result in higher effectiveness whereas in-process techniques are more flexible during their application as they need not be tailored to each task. However, the applicable area of in-process interventions is currently limited to handling tasks and worker notification.

Table 2.4 depicts the best achievable level of effectiveness for each risk factor for pre-process techniques and in-process techniques. Whereas force, as a risk factor, can be eliminated by both types of methods, posture and repetition lack such latitude.

To sum up, for the risk factor posture, there are either effective solutions or flexible ones. According to the literature review performed, a method combining both sets of advan-

tages has not been established as of yet. In the present work, this dilemma is referred to as the “effectiveness-flexibility trade-off”. The goal of this thesis was to bridge the trade-off by designing automation technology in a way such that it can eliminate or at least reduce the intensity of the postural load while still allowing workers to operate in-process.

Table 2.4: Maximum achievable effectiveness level for different risk factors. The cell colours correspond to the effectiveness levels depicted in Fig. 2.6.

	Pre-process	In-process
Posture	1 - eliminate	4 - behaviour
Force	1 - eliminate	1 - eliminate
Repetition	1 - eliminate	4 - behaviour

2.4 Conclusions

This chapter has provided an overview of how to tackle the problem of WMSDs at work. The workflow is comprised of two steps - ergonomic assessment to obtain an overview of the risk factors and ergonomic interventions, which represent concrete measures to diminish particular types of exposure. It is important to note that different risk factors, such as posture, force and repetition, have varying characteristics, requiring the development of tailored solutions.

Concerning ergonomic assessment methods, there are a variety of tools to choose from. Many of them are specialised for detecting particular aspects, while others provide an overall analysis. When selecting the appropriate method, it is important to consider whether the assessment has to be performed in an industrial or clinical/laboratory environment. Moreover, there is a trade-off between the detail of the results and the required investment. Finally, although previous research has reported evidence that assessments correlate to real WMSD risk [93, 34], the validity of each single criterion has not been fully proven to date.

Moreover, there are a variety of ergonomic intervention approaches for dealing with different risk factors. However, predominant solutions that match every situation have not been found. To make matters worse, the effectiveness-flexibility trade-off has been identified, which is based upon the fact that there are essentially either effective solutions or flexible solutions. An approach bestowed with both for the posture and repetition risk factors is not yet available.

3 Research Goal

In this chapter, the goal of the present thesis is stated. First, the scientific gap is identified, which was revealed through a literature research on the state of the art of WMSD prevention. To bridge this gap, a central hypothesis to be proven is formulated. The hypothesis is used as basis for defining the specific research questions. Finally, the research approach for each task and the structure of the upcoming chapters are described.

3.1 Research Gap

Chapter 2 provided an outline of the WMSD prevention strategies proposed. The most important lesson learned was that there is no dominant approach able to be applied to all problems. Each solution addressed different risk factors and differed in effectiveness, cost and flexibility. Unfortunately, effectiveness and flexibility have been demonstrated to be contradictory goals. Flexible solutions lack effectiveness and effective WMSD prevention demands a larger share of manual work from experts.

As an example, consider the effective workplace design measure: Ergonomists and workplace designers have to first gain knowledge on the underlying exposure of the task by using ergonomic assessment methods. The assessment involves analysing information recorded or observed from the task at hand. Thereafter, measures must be chosen and tailored to the workplace. These two tasks require a sizeable share of manual actions and expertise. The effort is multiplied with the amount of tasks to be (re-)designed. In flexible production scenarios, where product lines and tasks often change on a daily basis [94], ergonomic measures become the bottleneck and are thus neglected or unsatisfactorily implemented.

In-process prevention techniques are interesting as they are designed to be adaptable to changing situations and avoid tedious re-configuring of the workplace. However, their current area of application is limited to a few special cases. Moreover, as can be seen in Tab. 2.4, pre-process solutions outperform in-process solutions in terms of effectiveness by a wide margin. Especially for the posture and repetition risk factors, there is still a significant trade-off between effectiveness and flexibility. Only in case of force researchers and engineers have succeeded to resolve the effectiveness-flexibility trade-off.

Why has it been possible to realise an in-process measure addressing the risk factor, force, with high effectiveness? One answer is that the described solution possesses two essential components - a sensor system that can automatically detect the unfavourable situation and an actuator system which can automatically resolve the problem. A solution that misses one of these two components requires human involvement in order to function.

To reach an effective in-process system that addresses the posture risk factor, there is one piece of the puzzle missing - whereas there are promising approaches to detect critical postures, methods to achieve automated in-process workplace layout design are a definite research gap. Another issue is that in-process posture assessment, so far, primarily relies on IMUs. The drawback is that the sensors need to be worn and calibrated. In industrial environments, this is problematic. Worn devices limit freedom of movement. Every time a worker uses the system, it has to be prepared in advance, which includes a correct placement of the sensors and a potential calibration routine.

Ultimately, this thesis focuses on these two research gaps. In order to enable the realisation of an effective in-process solution, the topics of automated in-process workplace design and the image-based methods for in-process ergonomic assessment were addressed.

3.2 Research Hypotheses

Inspired by the concepts of human robot collaboration, the goal of this thesis was to develop a novel type of equipment capable of accomplishing WMSD prevention at effectiveness level 1 - 2 for the posture risk factor with the flexibility of an in-process method. Instead of IMUs, image-based methods were used to detect the critical posture. Thus, the method specifically avoids a tedious preparation and calibration process. Moreover, unlike inertial-based methods, image-based methods allow to additionally monitor the whole production environment including workplace and workpiece. On the other hand, the image-based approach possesses a variety of technological problems needing to be solved. It is challenging to design systems that accurately process images with strong variations in the appearance of subject and environment, illumination and other environmental parameters.

Nevertheless, such a technology can dramatically simplify the implementation of posture-related ergonomic interventions as time-consuming tasks are delegated to the computing system. In particular, enterprises operating flexible production lines can prevent their workers from posture-related WMSDs without the need for frequent workplace re-design. Overall, such a development would decrease the effectiveness gap between in-process and pre-process measures. To sum up, table 3.1 depicts the desired state of the art.

The goal of the present work can be summarised into the central research hypothesis:

*It is possible to realise an **in-process** WMSD prevention measure addressing the risk factor, **posture**, at an **intensity level of 1 - 2** on the effectiveness scale according to Goggins et al. by combining **image-based ergonomic assessment** with **in-process workplace layout design**.*

A system built to fulfil this hypothesis has to be experimentally verified (see Chapter 8). In particular, its in-process ability has to be verified by an execution time analysis of its components. The effectiveness has to be verified by applying the system on exemplary cases.

Table 3.1: Maximum achievable level of effectiveness after successful accomplishment of the research goal.

	Pre-process	In-process
Posture	1 - eliminate	1/2 - eliminate/intensity
Force	1 - eliminate	1 - eliminate
Repetition	1 - eliminate	4- behaviour

To exhibit the central hypothesis, there are multiple research questions that must be answered:

- **RQ1:** How does an in-process system look like which eliminates or reduces the exposure caused by awkward posture shortly after occurrence?
- **RQ2:** How can an automated image-based ergonomics assessment be realised?
- **RQ3:** How can an automated and adaptive in-process workplace design be realised on various types of actuators?

In the following, each research question is elaborated upon to outline which sub-tasks it implies.

RQ1: How does an in-process system look like which eliminates or reduces the exposure caused by awkward posture shortly after occurrence?

As the proposed solution represents a novel type of equipment, the first task is to create a vision of how such a device would look and operate. First, its tasks, use cases and requirements have to be identified. To implement such a device, the overall workflow has to be broken down into sub-problems that can then be technically solved. Moreover, a system able to react and initiate measures over a short time cannot be used when it poses hazards to the worker. Severe injuries have to be avoided at all costs. Hence, safety mechanisms must be integrated that prevent the worker from such threats. Additionally, these mechanisms have to be capable of compensating for human errors.

RQ2: How can an automated image-based ergonomics assessment be realised?

In order to appropriately react to the situation, it is essential that the ergonomic situation is appropriately assessed. Inertial sensor-based systems have their shortcomings in localisation accuracy. This leads to the problem that safety monitoring of the worker has to be realised by additional sensor systems. Moreover, before usage, these systems have to be calibrated to the user's anthropometry. Image-based methods, in contrast, can be applied to fulfil both aforementioned functions, which is why, in this thesis, the given task is approached by analysing the camera image. In terms of working postures, the posture the worker adopts must be recognised and evaluated. The recognition method needs to be robust to variations in human anthropometrics and environmental conditions.

RQ3: How can an automated and adaptive in-process workplace design be realised on various types of actuators?

Most level 1 - 2 measures addressing the posture risk factor require careful manual workplace planning and design. Therefore, to create an in-process measure requires automating this designing process. The process of automatically creating solutions has to be fast enough to provide immediate feedback to the worker. The focus lies on providing the required data to manifest the feedback, but not the feedback itself. The task of designing an appropriate presentation of the information to the human being is outside of the scope of this thesis. To adapt a workplace layout, manipulators or actuators have to be employed. There are different types of actuators, such as height-adjustable platforms, robot manipulators or tilting platforms. The developed method has to arrive at acceptable solutions with different available Degrees of Freedom (DoF).

3.3 Research Approach

The upcoming chapters derive answers to the research questions posed. Chapter 4 provides an answer to RQ1. In detail, use cases and requirements are developed from RQ1. Afterwards, the system concept is presented identifying the technological components to be implemented. Chapter 5 concentrates on finding an answer to research question, RQ2. The posture assessment algorithm is a core component of the WPC system. The developed algorithm is put forth, including the design decisions made. Finally, Chapter 6 presents an answer to RQ3. The chapter documents the algorithmic framework developed for the WPC to optimise workplace layout. This framework is then adapted to operate with various actuators. Finally, the dynamic behaviour of such a system is analysed.

The concept chapters are structured in a standardised manner containing four parts: "problem statement", "design alternatives", the solution itself and "conclusions". The section, "problem statement" reformulates the research question to be investigated. Rather than repeating it, there is an attempt to elaborate upon and transform the given

task into known scientific and engineering problems. Thus, the problems can be solved with existing engineering methods. The "design alternatives" sections summarise the related work to the transformed problem outlining design decisions made. Furthermore, the criteria for how the design alternatives for the WPC were chosen are described in the sub-sections dubbed "selected approach". Although the method selection was based on a detailed discussion among available alternatives, it does not mean that it is the perfect choice. Sub-section "risk analysis" attempts to identify potential scientific risks when realising the elected methods. The following sections present the concept of the specific solution. The structure for this part is not standardised, but rather depends on the content of the presented solution. Finally, each chapter ends with a "conclusions" section where the most relevant outcomes are summarised.

4 The Working Posture Controller

This chapter describes the concept underlying the novel type of equipment envisioned in the research goal chapter 3. This device is referred to as the "Working Posture Controller" (WPC). The task of the WPC is to enable workers to adopt neutral working postures while performing their tasks at hand. Situations where the worker adopts an awkward posture for a long time are to be avoided.

4.1 Problem Statement

This chapter attempts to address research question, RQ1:

How does an in-process system look like which eliminates or reduces the exposure caused by awkward posture shortly after occurrence?

To determine the sub-tasks, the question is analysed word by word:

Firstly, the term "in-process" indicates that the system must detect the ergonomically unfavourable situation shortly after it happens. In this case, the unfavourable situation is an awkward posture adopted by the worker. By "shortly", a maximum duration of a few seconds is meant. Secondly, the words, "eliminate" and "reduce exposure", imply that the WPC has to actively intervene. Being passive and leaving the intervention to the worker would not result in a level 1 - 2 measure (see Sec. 2.3 for definition). Additionally, as the system actively intervenes, it has to be ensured that the intervention is safe for the human being. Thirdly, the word "How" suggests the question concerns the operation principle and the workflow of the system. To illustrate this principle and to identify requirements, potential use cases of the WPC are derived. Figure 4.1 summarises the research questions to be answered in this chapter.

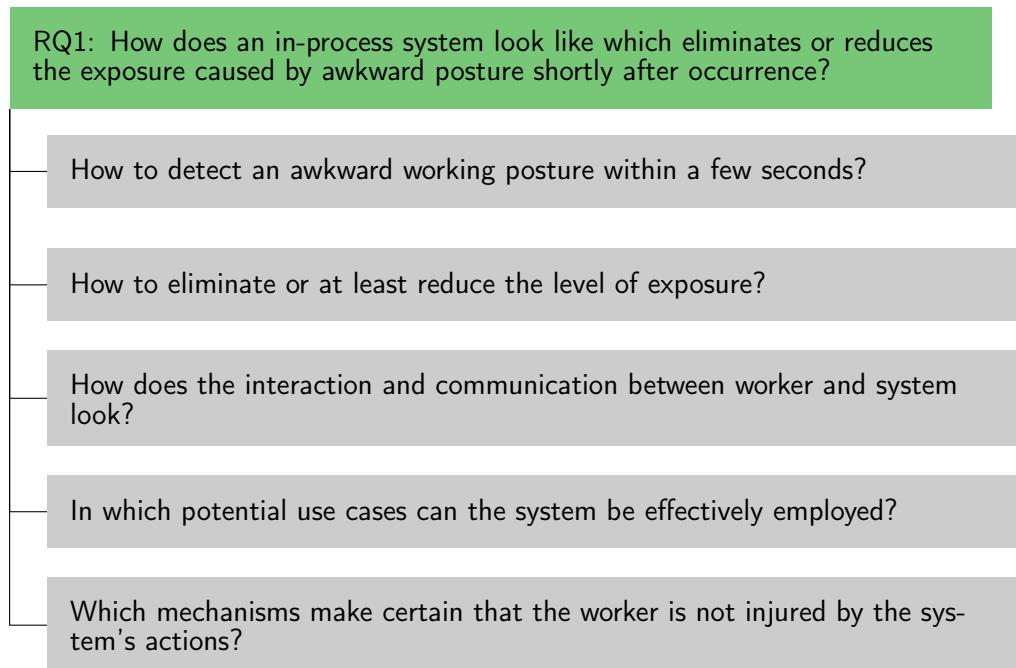


Figure 4.1: Summary of the sub-research questions RQ1 implies.

4.2 Design Alternatives

4.2.1 Detecting an Awkward Posture

The task of detecting an awkward posture can be accomplished by applying ergonomic assessment methods that evaluate the present posture. Li and Buckle [3] classified ergonomic assessment tools into the categories self-reports, direct methods and observational methods. The tools mentioned and discussed in Sec. 2.2 represent the available alternatives.

Selected Approach

First of all, self-reports as design option for the WPC can be excluded. Self-reports necessitate the worker to recognise the posture by themselves. If the WPC awkward posture detection merely relied on the estimation of the worker, it would be a level 4 intervention. The remaining two methods are both suitable for use in the WPC system. Direct methods have the advantage that the relevant variables can be automatically measured. However, they often require a calibration routine. In addition, the user must wear bulky equipment. More complex observational methods provide a detailed analysis of the posture and do not limit freedom of movement. Yet, they currently need trained personnel to properly use the tools.

In this work, direct methods were combined with observational methods. The awkward posture was detected by the observation of the process. However, a direct method was

used to acquire the data necessary to automatically fill out the worksheet. Among the various observational assessment tools, the static posture assessment section of the “ergonomics assessment worksheet” (EAWS) [34] was chosen. The EAWS is a method that provides numeric risk scores. As described in Sec. 2.2, a numeric assessment result is imperative for performing effective workplace design as it is sufficiently nuanced to depict the level of success of an intervention. The EAWS is one of the most detailed expert-screening methods [4]. Further, the creators of the EAWS stated that it was designed to be compliant with other internationally accepted posture assessment methods, such as RULA [28] and OWAS [30].

Unlike the alternatives RULA and REBA, the EAWS applies the work-sampling scheme containing additional information about the duration of postures adopted. Work-sampling is vital for continuous ergonomic monitoring, since it avoids that the system frequently intervenes even when the duration of the awkward posture is only marginal. Section 6.8 discusses this issue in detail.

Alternatively, OWAS or the PATH method [32] could have been chosen. The EAWS posture section is nearly identical to OWAS. However, EAWS encompasses further ergonomic aspects, such as action forces, material handling and repetition. A postural analysis using the EAWS, therefore, could complement other assessments. PATH is more nuanced than EAWS and OWAS as it even considers neck postures. However, it does not provide a decision rule when to intervene. Thus, it is more suitable for tasks where the goal is to compare different workplace designs.

Table 4.1 summarises the criteria that were considered for the selection of the awkward posture detection approach.

Table 4.1: Overview of awkward posture detection alternatives. Columns highlighted in green represent the selected approach.

Criterion	Self-reports	Observational methods	Direct methods
Can be automated	no	yes	yes
Exactness [2]	low	medium	high
Bulky equipment required	no	no	yes
Calibration required	no	no	yes

Risk Analysis

The EAWS belongs to the group of observational ergonomics assessment tools. Although it is one of the most elaborate members of its group, the ability of observational methods to accurately predict the WMSD risks is controversial. Schaub et al. [34] state that the EAWS agrees to important European guidelines and that its practical relevance has been tested using ergonomists as ground truth. Nevertheless, no large scale study has been found to statistically prove its validity. Furthermore, automatising the EAWS requires

developing a system, which is able to detect 16 postures excluding the asymmetry effects (see Sec. A of the Appendix). This task poses a remarkable challenge for the automated posture assessment component.

4.2.2 Measures against Postural Load

The literature features three main options for avoiding postural load: working differently (job rotation or pausing), changing human behaviour (training or warning systems) or modifying the working environment (workplace design). A summary of these methods was provided in Sec. 2.3. An outline of the available options is provided in Figure 4.2.

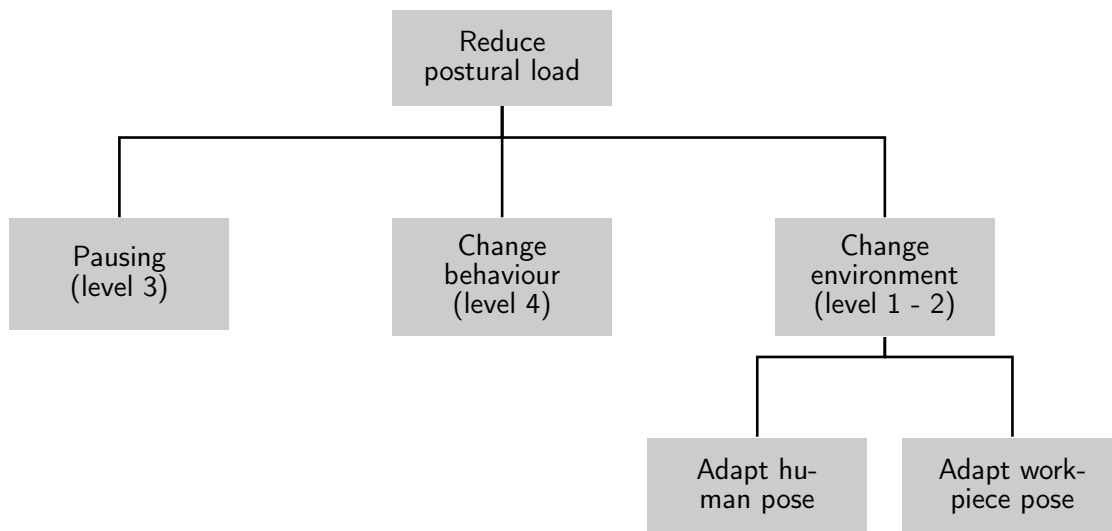


Figure 4.2: Measures for avoiding awkward postures.

Selected Approach

While job rotations or pauses are organisational measures that, at most, achieve level 3 on the effectiveness scale, measures relying on a change in human behaviour are rated as level 4. Only altering the work environment is able to achieve an improvement at level 1 - 2. In terms of postural load, it implies refining the spatial relationship between human and the object to be processed, e.g., the workpiece. This change can be realised in two ways - either relocate the worker or the workpiece. An example for the first case would be the worker standing on a movable platform. While altering height or distance to the workpiece can be realised, problems arise as the actuator cannot safely tilt a human. Hence, the degrees of freedom of such a system are limited. The other alternative is changing the pose of the workpiece, which offers more freedom of movement. Inspired by the work of Busch [89], adapting the workpiece pose was chosen as approach. However, the goal is to realise the motion planning of the actuator in-process instead of in the design stage. An illustration of this can be found in Fig. 4.3. The left panel exhibits a non-optimal pose of the workpiece forcing the human to adopt an awkward posture

when working on it. The right panel portrays the configuration that leads to a neutral working posture. Table 4.2 outlines the criteria for the selected WPC measure to reduce the postural load.

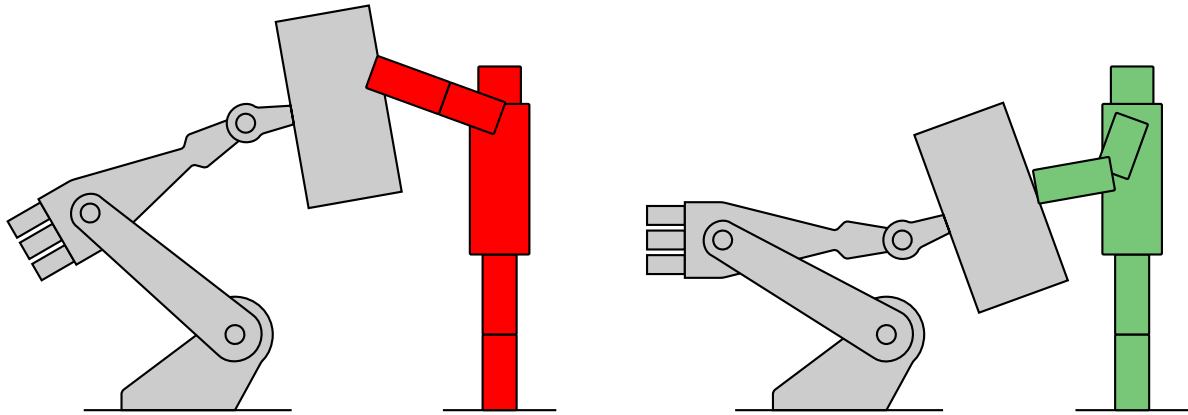


Figure 4.3: An actuator holding the workpiece is used to adjust the workplace. The left panel illustrates an ergonomically unfavourable actuator configuration, which is corrected in the right panel.

Table 4.2: Overview of awkward posture avoidance alternatives. Columns highlighted in green represent the selected approach.

Criterion	Break	Change behaviour	Adapt human pose	Adapt work-piece pose
Effectiveness level	3	4	1- 2	1 - 2
Realisable DoF	none	none	relocation	relocation and tilting

Risk Analysis

Adapting the workpiece pose has been elected as the WPC measure to avoid or diminish postural load. In order to realise a workpiece pose adjustment, an actuator is required, which is powerful enough to carry heavy workpieces. However, such powerful actuators pose safety threads which have to be carefully controlled. Furthermore, one challenge in realising the WPC is actually proving its financial justification. Using robot actuators implies investing a five to six digit amount of Euros, which then has to be compensated by an at least equally powerful benefit. One solution could be using height-adjustable actuators available for a lower price. Unfortunately, such a type of actuator only supports a limited range of adjustments.

4.3 System Concept of the WPC

Section 4.1 concluded that the WPC requires a system to detect awkward postures as well as one to steer clear of remaining in awkward working postures. In the following, the former is referred to as the posture assessment component and the latter as the posture optimisation component. The following approach from Sec. 4.2 has been derived - to perceive the awkward posture, a combination of observational and direct methods were chosen. If level 1 - 2 interventions are targeted, the equipment must automatically manipulate the environment. A manipulation of the workpiece pose is preferred based on more opportunities for adjusting the workpiece pose. Additionally, manipulators with more degrees of freedom can be applied. In this section, the workflow is described at a higher level. The specific algorithms to reach automated posture assessment and posture optimisation will be addressed in the chapters 5 and 6, respectively.

The workflow of the WPC is designed as follows (see Fig. 4.4 for visualisation) - the posture assessment component continuously monitors the posture of the worker and yields a posture-related risk score according to a modified version of the EAWS [34]. Posture optimisation comes into play when the worker is forced to adopt awkward postures in order to accomplish the task. A user has to adopt an awkward posture when the targeted region on the workpiece to be processed is hardly visible or not accessible. As a consequence, the spatial relationship between the worker and workpiece has to be modified such that the task requirements can be met even when adopting the desired posture. To find this adjustment, the task at hand needs to be considered. The WPC must know which region on the workpiece the worker intends to interact with. Furthermore, it has to recognise the direction of the working movement. These parameters were extracted from the worker's current posture. Localising the hands permits targeting the region whereas estimating forearm orientation yields an approximation of the working motion direction. The modification of the spatial relationship between worker and workpiece is executed by a manipulator that holds the part. Depending on the available type of actuator, such as an industrial robot, a height-adjustable platform or a tilting table, the posture optimisation considers the available DoF and determines the best solution within these limitations.

Subsequently, the loop starts anew - the posture assessment component observes the new behaviour of the worker. When the current environmental parameters do not produce an acceptable posture assessment result, the WPC initiates a new adjustment. As this behaviour resembles a control loop, the name of this system has been chosen accordingly.

4.4 Use Cases

After having described the functional principle of the WPC, scenarios were derived where this type of equipment could help reducing postural load. Here, four possible use cases are presented, but no claim of completeness is raised.

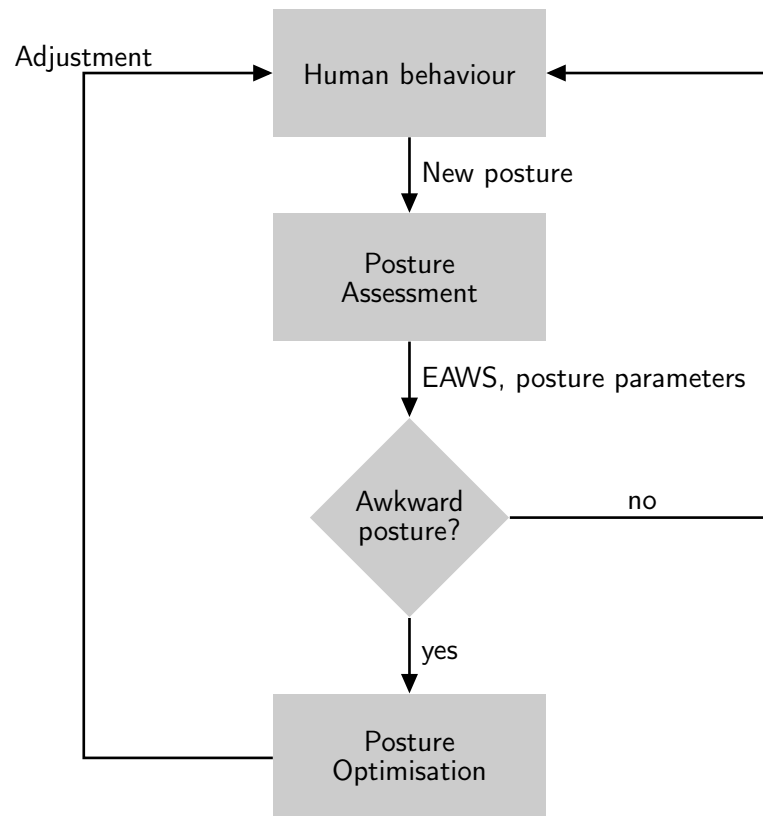


Figure 4.4: The basic WPC control loop.

Avoidance of Forced Posture

When workpieces to be processed are not in range or not visible for inspection, the worker is forced to adopt awkward postures. Especially in the case of large products, this serves as a threat because actions have to be likely performed at different sides and parts. For example, when welding pipes, the torch held by the worker has to move around the workpieces [89] resulting in a constant change of location and working direction. Moreover, a worker modifying electric engines of dimensions 1m x 1m x 1m might need to first disassemble parts from the top. Afterwards, repair operations might require access via the bottom of the engine. Every time the worker has to process a part at a distant location for a longer time, the WPC would change the pose of the workpiece to facilitate the worker accomplishing the job in a neutral posture.

Matching the Working Height and Posture to the Type of Task

According to Schmidtke and Jastrzebska-Fraczek [95], the effective working heights for a standing worker are based on the type of task being engaged in. For precision work, including visual inspection and quick motions, a working height between 100cm – 120cm is recommended. Work containing forceful motions is best fulfilled at a height between 90cm – 110cm and heavy work containing whole-body movements suits heights between 70cm – 90cm. The exact value depends on the worker's anthropometry. Note that these recommendations assume working with products of smaller dimensions. If the product is

bigger, the range of recommended working heights becomes broader. If the system knows the type of task and product size, it can incorporate this information into the posture optimisation algorithm.

Enable a Convenient Direction of Force Exertion

According to Schaub et al. [44], the maximum acceptable force exposure depends on the worker's posture and the direction the force is being applied in. For instance, when standing upright, larger force exertions are acceptable compared to bent postures. Additionally, higher maximum forces can be tolerated if the force direction points downwards rather than upwards. A task that originally requires an upward force direction can be transformed into one requiring a downward direction simply by flipping the workpiece.

Initiating a Dynamic Postural Change

Although upright standing postures and upright sitting postures are the most favourable according to the EAWS [34], remaining in one posture for the whole working day causes monotonous strain. Hence, it is essential that the worker changes posture from time to time. As commendable guideline, Neuhaus et al. [51] encouraged sitting 50% of the time and standing upright for the other half. The WPC could automatically identify the current posture and measure the time the worker has remained in it. If the time exceeds a threshold, the WPC changes the workpiece pose, which in turn allows the worker to alter his or her posture.

4.5 Safety Mechanisms

The actuators of the WPC share a common workspace with the worker. Humans and systems are coupled to each other through the workpiece. Specifically robot manipulators are able to generate high forces, torques and speeds. This creates lethal threats in the case of clamping or when the actuator collides with the human. Haddadin et al. [96] mentioned the following possible contact scenarios to consider - unconstrained impacts, partially constrained impacts and resulting secondary impacts. Thereby, constrained impacts, such as clamping in robot structures or walls, are highly dangerous and possibly result in life-endangering situations. Another safety issue to be considered is when a robot holds the workpiece such that certain parts are not completely fixed. If the object is tilted, these parts could potentially fall, leading to damage or injuries.

Taking into account these hazards, integrating safety mechanism is mandatory. In the following, three mechanisms that can enhance the safety of the WPC when interacting with humans are briefly described. Fully implementing all mechanisms would significantly slow down the workflow. However, the focus of this thesis was not on developing a time and safety optimised concept, but instead to ensure basic safety. The selection of which

mechanism to include in the final system depends on the specific use case. The safety mechanisms are derived from robot and machine safety standards:

- DIN EN ISO 10218-1 [97]
- DIN EN 60204-1:2014-10 [98]
- ISO/TS 15066 [99]

The following sections further elaborate upon the mechanisms. Figure 4.5 summarises the presented mechanisms and which threat they address.

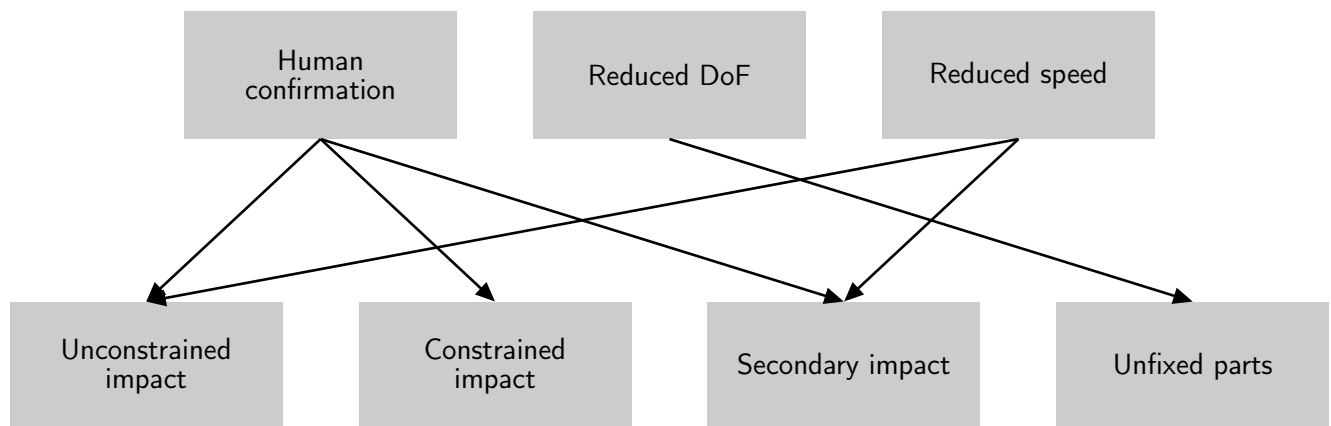


Figure 4.5: Summary of safety mechanisms and the threats they address.

Human Confirmation for Adjustments

Unexpected behaviour that brings about injuries can be dramatically reduced when every action of the system must be confirmed by the human worker. A movement of the actuator while the worker is nearby can lead to confusion or a contact scenario in the worst case. A confirmation mechanism makes certain that the worker at least expects the motion of the actuator. Moreover, an alternative manual operation mode of the system can be instituted where the user completely controls the actuator. In any case, the control interface is recommended to be placed outside of the manipulator's range so as to eliminate contact risk. The actuator has to cease its motion as soon as the human enters the workplace, and there is no collaborative workspace. Moreover, the standard DIN EN ISO 10218-1 [97] requires the confirmation mechanism to work according to a three-phase confirmation regimen. In situations where the confirmation is unintended and a human being is in a shock situation, they can end the motion by either releasing the switch or pushing it further. Figure 4.6 illustrates the human confirmation mechanism integrated into the original WPC workflow.

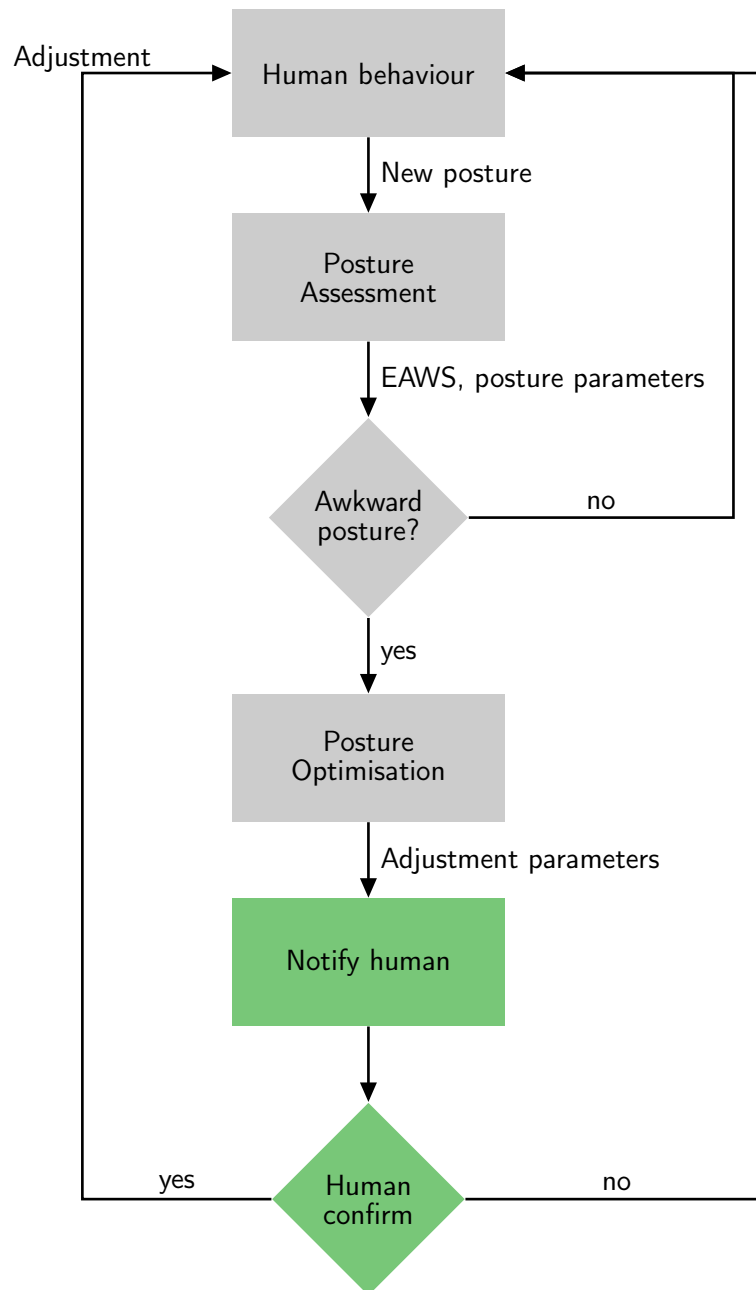


Figure 4.6: Safety enhanced WPC workflow.

Reduced Degrees of Freedom

Limiting the DoF reduces the range of the actuator and renders motions more predictable. If the human being stands in front of the system, the chance of a collision is markedly reduced when the actuator is only permitted to move vertically. By restricting the global orientation of the workpiece, this mechanism can also prevent unfixed parts from falling out. The limitation of the end effector orientation can be incorporated into the hardware and the posture optimisation algorithm.

Reduced Speed

In the case the rather conservative confirmation mechanism unacceptably reduces the efficiency of the process, the WPC system can be designed as a collaborative system where the actuator initiates an adjustment although the user is located in the collaborative space. In this design, the system must fulfil the rules defined in ISO/TS 15066 [99]. This standard defines required protective measures and stopping functions for collaborative workspaces.

One crucial control for the safety of the interaction within the collaborative workspace is the speed of the actuator. According to ISO 10218-1 [97], the maximum allowed speed of the end effector in physical cooperation scenarios is 250 mm/s. This rather conservative threshold has been selected so the worker may react in cases of an unexpected motion taking place [96]. In situations of contact, the reduced speed results in a lesser force released onto the worker. ISO/TS 15066 [99] attempts to provide a more nuanced scheme. A formula is supplied to calculate the permissible speed depending on the protective separation distance and the distance of the actuator to the human being. Furthermore, maximum tolerable threshold values for contact situations depending on force, pressure, contact surface and part of the body are defined.

4.6 Conclusions

In this chapter, the overall concept of the WPC has been put forth. Ultimately, the idea is to have a system with an automated ergonomics analysis module and an automated posture optimisation module, which adapts the workplace layout. To foster the automated ergonomic assessment, the manual assessment process was taken as a model and the EAWS was chosen as the reference tool. The automated workplace layout design was then realised through an actuator holding the workpiece. Both components require to be computationally efficient so feedback can be provided within a few seconds.

The interaction workflow between a user and the system was then presented. Derived from this high-level system concept, use case scenarios for the WPC were provided that were mostly comprised of matching the posture to the process and enabling the worker to better apply force. As the worker and the system's actuator share a workspace, the WPC has to fulfil certain safety criteria. To reach this goal, safety mechanisms were proposed. One of the most effective ways, in general, is to integrate a human confirmation mechanism where the worker is outside of the danger area. Decreasing the degrees of freedom and end effector speed further diminishes the risk of injuries.

5 Posture Assessment

This chapter explains the algorithms developed to formulate the posture assessment component of the WPC. The focus was on developing a method to estimate the location of the joint angles, which are vital for recognising the posture. Having identified those postures in a process, the standard EAWS scheme can be applied to obtain a numeric risk score.

5.1 Problem Statement

This chapter addresses the research question RQ2:

How can an automated image-based ergonomics assessment be realised?

As Chapter 4.2 stated, one core component of the WPC was an algorithm able to provide a numeric score representing the postural risk. In order to initiate an in-process intervention, the assessment must be able to continuously determine the ergonomic risk scores over short time intervals. Due to its numeric result format and the work-sampling operation principle, the EAWS [34] was selected as the method to be automated (see Sec. 4.2.1 for a detailed discussion). As a side condition, Chapter 3 has put forth that an image-based approach was selected. A manual assessment according to the EAWS requires the ergonomist to determine which postures a subject has adopted and for how long. Although the EAWS does not contain all feasible postures of the human body, the present use case assumes that all the observed postures in the process can be assigned an EAWS posture. An automated assessment would simulate the manual process by automatically determining over short durations which posture the recorded image depicts. The recognised postures and their time share since the beginning of the recording can then be translated into intermediate posture scores. The sum of the intermediate posture scores yields the current risk score. An example assessment is performed in Sec. A of the Appendix.

The task of automatically recognising a particular posture from a given finite set of options is known in machine learning literature as classification. Classification is characterised as the task of establishing which pre-defined category an observation belongs to. The decision can either be made on the basis of a set of manually developed rules (as in the case of manual EAWS assessment) or by an algorithm learning the decision pattern

from a set of training observations. In the training data, the corresponding categories are known. The training data-based approach was centred around as the given observations were images. Decision rules derived from the raw input pixels were expected to be highly complex and hard to manually formulate. For the posture classification problem, the observations correspond to the images acquired by the image sensor. The categories are the postures to be identified. Based on the fact that there were more than two postures in the EAWS, this is a multi-class classification problem. Normally, classifying raw image data does not produce robust results. Rather, the key to high accuracy is to find features which invariantly and uniquely represent a posture category [100]. The features described in this chapter require a sophisticated method to extract, which is why a large proportion of this chapter will deal with the actual obtaining of them, including the camera setup and the algorithm used.

Figure 5.1 summarises the research questions to be answered in this chapter.

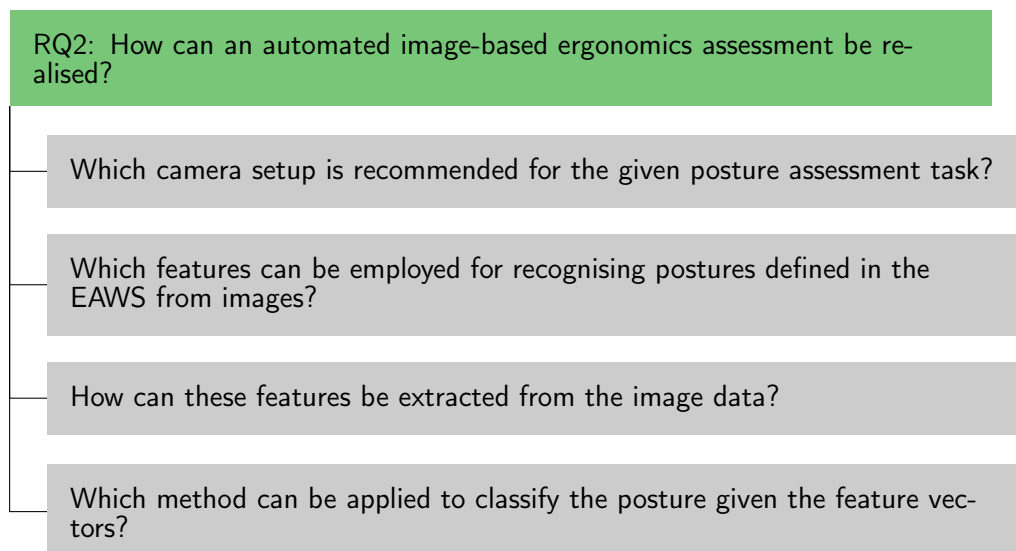


Figure 5.1: Summary of sub-research questions RQ2 implies.

5.2 Design Alternatives

5.2.1 Posture Classification

Classification approaches have been proposed to recognise exemplary postures, such as "standing", "bending", "sitting", "lying", "crouching" or "squatting" [101, 102, 103, 104, 105, 106, 107]. Certain posture classifiers have been developed in the context of a specific application area. For instance, there are works on home care [102, 103] or the translation of hand sign language [108, 109]. Strongly related to posture classification is the task of action recognition, where a motion is labelled as an action category. In this case, methods are used that consider a combination of postures. Action recognition

approaches are included in this overview because the features extracted are similar to those used in posture classification. In order to realise a posture classifier, two components must be selected - the classifier and the features. A variety of classifiers have been utilised: neural networks and their variants [110, 111, 102, 103], support vector machines [112, 105, 113], random forests [109], nearest neighbour-based classifiers with various distance measures [114, 106], threshold-based classifiers [104], linear classifiers [108] and Naïve Bayes [115, 116]. Conversely, in the field of human action recognition, which is basically a classification task over a sequence of postures, hidden Markov models are the favoured technique [101, 117, 118, 107, 119, 120].

Similar to classification algorithms, a variety of different image features for posture classification have been proposed. Figure 5.2 summarises the existing approaches. Briefly, the posture classification features can be categorised into two groups - features that are directly computed from the image (direct features) and features computed from intermediate features, including locations of the parts of the body (indirect features).

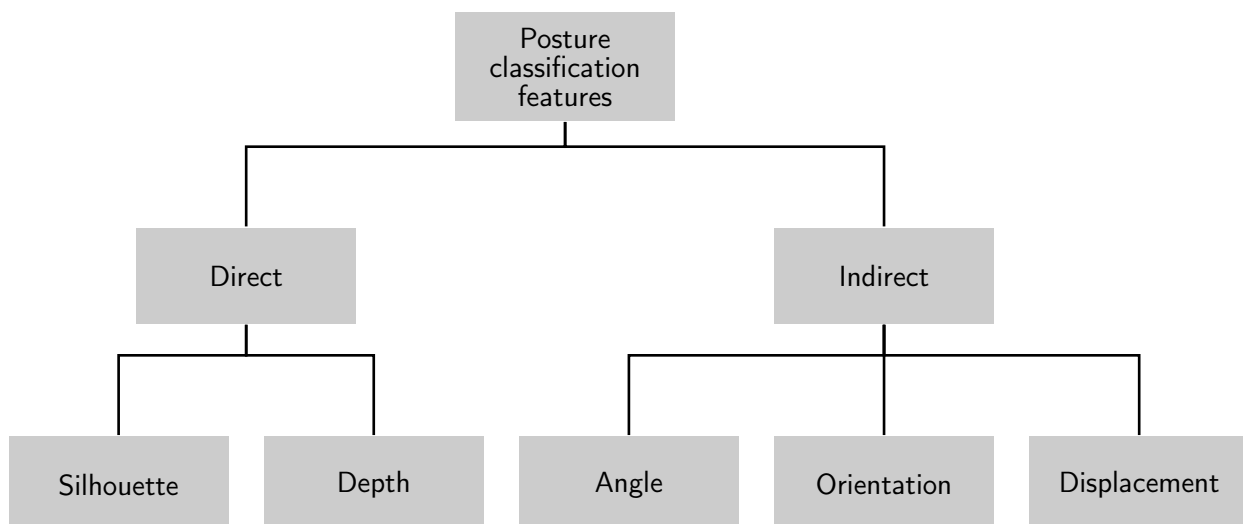


Figure 5.2: Overview of feature types used for posture classification.

Direct Features

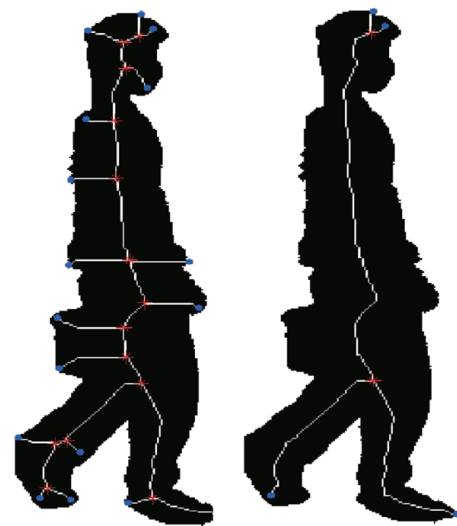
Most direct features are computed from silhouette images [110, 101, 117, 111, 102, 103] or depth images [104, 118, 109]. Zerrouki and Houacine [103] and Just et al. [108] used Red Green Blue (RGB) colour images. Over the period of the development of the posture assessment component (2012 - 2014), RGB images as input modalities were not as popular as depth or silhouette images as the appearance of the subject varied depending on their clothes and environmental illumination. Only recent developments, such as in Taylor et al. [121], suggested methods relying on RGB images, which achieve promising results. For this reason, RGB features were excluded as alternatives.

Silhouettes (see Figure 5.3a for an example) represent the segmented person or object in the image and can be procured by background subtraction or another segmentation algo-

rithm. Silhouettes are signified by binary images where a pixel value indicates whether the pixel belongs to the foreground or background. Rather than isolated pixels, a foreground normally comprises connected regions. Thus, most features attempt to describe the shape of such regions. A popular approach for capturing the shape is probabilistic projection mapping [101, 117, 102]. The silhouette is projected onto the principal axes. The resulting 1D profiles represent the minimum or maximum foreground extent in each row or column of the image. Line-wise histograms were then created to express the variance between silhouettes of the same posture class. As a result of silhouette edges possibly being non-smooth and noisy, an accepted post-processing step is to apply a Fourier transformation onto the profiles and filter out the high-frequency components of the signal [102]. A different approach to creating features from silhouettes are shock graphs (see Figure 5.3b), just as used in Hussain et al. [111]. Shock graphs [122] are produced by the skeletonisation of the silhouette. A skeletonisation algorithm successively applies morphological erosion operations onto the image until only tree-like structures of the width of one pixel remain. Using shock graphs, concave and convex structures can be detected.



(a) Silhouette image



(b) Shock graph by Hussain et al. [111]

Figure 5.3: Examples of direct features for posture classification.

Processing depth images or depth maps has recently become popular because of the release of low-priced Red Green Blue - Depth (RGBD) sensors, such as the Microsoft Kinect®. The value of a depth map pixel denotes the distance along a ray from the pixel on the camera sensor to the first obstacle the ray collides with. Therefore, depth images contain 2.5D information. Additionally, background subtraction within depth images constructs a silhouette image. This processing technique is a popular choice [11, 123] as depth images vary less under illumination changes compared to RGB images. Thus, depth images possess 2.5D information as well as silhouette information, allowing the use of the techniques mentioned before. Employing 2.5D information, Diraco et al. [104]

created a Reeb graph. This structure can be seen as the shock graph derived from 2.5D data. Kuznetsova et al. [109] put forward computing ESF descriptors, a statistical approach to describing the data. The depth image was first masked out by the silhouette. Subsequently, random pairs and triplets of points were sampled from the region and their 3D distance, angular and area measurements computed. The histogram representing the distribution of these computed measurements was used as the feature vector. Another approach is to compute a histogram of oriented gradient (HOG) [124] on depth data [119]. HOGs are state of the art features in RGB images, e.g., for the localisation of human beings.

Indirect Features

The methods described earlier directly extract the features for posture classification from the image. However, many approaches [104, 112, 114, 105, 106, 107, 125, 126, 119, 113, 120, 115, 116] prefer indirect features. First, distinctive parts of the body, in particular joints, are localised. Afterwards, features are computed from combinations of the identified positions. Exemplary joints are "centre of the left hand", "left elbow" or "left ankle". The joint positions are described by 3D coordinates. A problem is that joint coordinates indicative of the same posture vary depending on the position of the subject, the size of the body segments and the global orientation of the body. Features computed from joint coordinates attempt transformation of the position coordinates into an invariant representation. This invariance can be achieved with joint angles [112, 106, 105, 125, 107]. Joint angles are the angles computed between two neighbouring body segments. Their advantage is that they can be easily interpreted by human beings. Incorporating prior knowledge, such as typical angles for specific postures or debugging the classification, is simplified. Joint angles are invariant to body segment sizes as the same angle between different segment lengths produces the identical feature vector. Unfortunately, "unwanted invariances" are obtained. Joint angle features can be ambiguous, as three 3D position vectors (nine-dimensional space) are mapped into one scalar value (one-dimensional space). This potentially results in distinct postures represented by the same angular values (see Fig. 5.4 for an example).

Joint or segment orientations [119, 120] are another invariant representation. The idea is to compute the normalised 3D direction vector of each body segment. Whereas joint angles signify the relative rotation between two neighbouring segments, joint orientations denote the rotation in the global coordinate system. By transforming the orientation vectors into polar coordinates [120], one dimension per segment can be discarded. Another way to introduce invariance is by using relative position displacements. Reyes et al. [126] set a fixed reference point on the skeleton and expressed all other joint coordinates relative to this reference. Alternatively, rather than fixing the reference point, the feature vector comprises pairwise distances between joints [113, 115, 116]. As the full feature space of pairwise joint combinations is high, Yang and Tian [115, 116] thought to apply a principal component analysis on the feature space by only selecting the pairs with the highest variance. When dealing with relative joint coordinates, it is important to

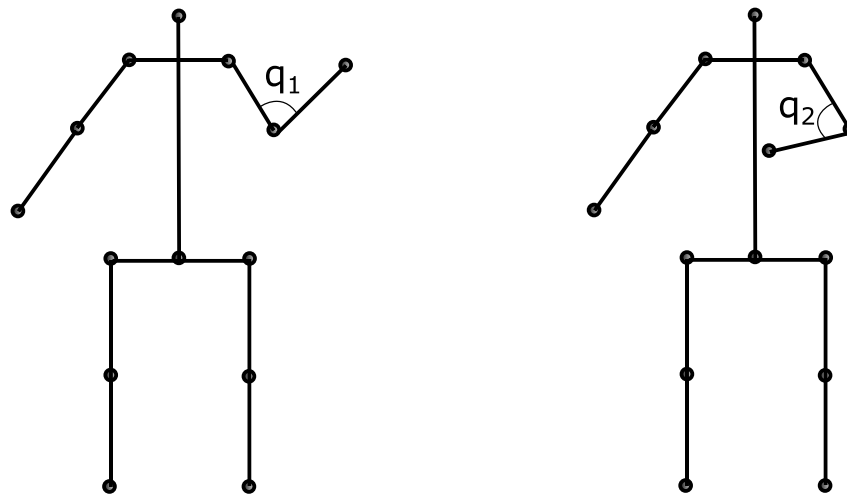


Figure 5.4: Ambiguity of joint angle features. Both postures are represented by the same joint angle vector as the joint angle values q_1 and q_2 are similar.

normalise them, otherwise, subjects with varied body sizes would result in different features. However, normalising does not guarantee invariance to different ratios of segment sizes, e.g., in the case of exceptionally longer arms or legs. Another approach is to take the plane orientation from triplets of joints, just as performed by Yao et al. [127]. Three points are needed to construct a plane and the orientation of the plane is represented by its normal. Using joint displacements, human expertise can be incorporated into the features by pre-classifying them into categories, such as "hand is near X" or "hand is above X" [119], where X represents a particular joint.

Selected Approach

Summarising the overview on posture classification features, direct features, in contrast to the indirect ones, have the advantage that they do not require an intermediate joint localisation step. However, methods from this family cannot provide body size and view point invariance. As joint positions have to be determined for the posture optimisation method (see Chapter 6), anyways, the indirect feature approach was followed. Using these types of features enables reusing the computed joint positions.

Among the indirect features, there was the choice between joint angles, joint orientations and displacement features. Joint angle features, versus displacement features, have the advantage of being invariant to body size and viewpoint. In comparison to orientation features, they represent similar information. Finally, joint angle features were implemented. In 3D, the joint angles were combined with orientation features to resolve the ambiguity of two postures having the same angular representation. In 2D, orientation features were not used, as a particular parametrisation scheme can avoid ambiguities. Table 5.1 summarises the criteria considered when selecting features for the posture classification component.

Table 5.1: Overview of alternative features for posture classification. Columns highlighted in green represent the selected approach.

Criterion	Silhouette	Depth	Angle	Orientation	Displacement
Compute joint positions	no	no	yes	yes	yes
Body size invariance	no	no	yes	yes	no
View point invariance	no	no	yes	yes	no
Ambiguous representation	no	no	yes	no	no
Parameters per joint			1	2 (3D)	3 (3D)

Risk Analysis

Indirect features heavily depend on the robustness and accuracy of the joint localisation algorithm used, which is still an open problem in the field of computer vision. If the correctness of the computed joint positions cannot be ensured, postures cannot be reliably classified. Furthermore, although the EAWS recommends a human practitioner to use joint angles and locations to distinguish the defined postures, it does not mean that these features themselves are sufficient to correctly classify each posture. A trained ergonomist might internally recognise the stance based on additional features, which are not explicitly expressed in the guidelines.

5.2.2 Joint Localisation

As stated before, the indirect posture classification approach requires a routine to identify the 3D positions of the joints of interest. In context of this thesis, this task is referred to as joint localisation. The biggest challenge in joint localisation is variance - identical postures can be depicted by countless different input images. A localisation algorithm has to localise the body parts under varying circumstances. Poppe [128] mentioned that motion and appearance of the subject can strongly vary. Additionally, the algorithm has to operate under unique environmental conditions, such as varying camera viewpoint, illumination and occluded view. The general approach to dealing with this variance is to introduce assumptions regarding subject and environment. A subset of conditions are fixed to reduce the overall variance. Moeslund and Granum [129] provided a list of the most common assumptions ranked according to their prevalence. There are extensive surveys [129, 128, 130] giving a detailed overview on joint localisation methods. In general, the approaches proposed in the literature can be categorised into three groups [131]: discriminative approaches, filtering approaches and generative methods. In the following, the main idea of every approach is outlined. The input of every method is a silhouette, a depth or an RGB image. In the following, the notation I is used to represent the image. The goal is to determine the joint locations and joint angles, which will be referred to as model parameters. The model parameters are denoted by the symbol \mathbf{q} .

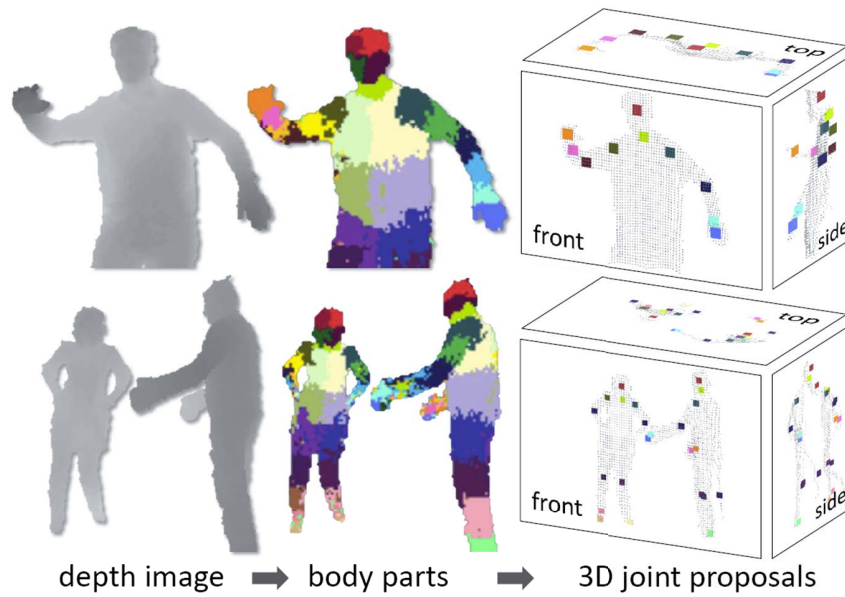


Figure 5.5: Shotton et al. [11] used intermediate body part maps to compute joint coordinates.

Discriminative Approaches

Discriminative approaches attempt to learn a mapping $\mathbf{q}(I)$ between an image I of the image space and model parameters \mathbf{q} of the parameter space from a set of examples $\{(I_1, \mathbf{q}_1), \dots, (I_n, \mathbf{q}_n)\}$. As image space, as well as parameter space, can be high dimensional, a large amount of densely sampled data is required. Discriminative methods can be divided into two groups - classification [11, 132, 133, 134] and regression approaches [135, 136, 137]. Regression methods strive to directly learn the mapping between image and parameters whereas classification methods learn discrete high-level features from the data and employ these features in computing the target model parameters. One example of a classification approach is featured in Shotton et al. [11] - they learned a mapping between a single pixel in the image and the part of the body it belonged to. After classification of each pixel, cluster centres in the resulting image represented the model parameters (see fig. 5.5).

The greatest advantage of discriminative methods is that, once the mapping has been learned, they can perform quickly and are able to deal with occlusions if the corresponding examples are included into the data set. However, the greatest challenge with them lies in developing methods for generating a massive amount of labelled data that contains all the necessary variation. As an example, Shotton et al. [11] produced around one million depth images containing subjects of various anthropometry. Each image contained annotations where each pixel was assigned the part of the body it belonged to. Manually creating such a data set is simply not feasible. Automatically creating this sort of data set requires computer graphics tools for human modelling. If there are postures that significantly differ from the examples utilised for training, the localisation can be inaccurate or

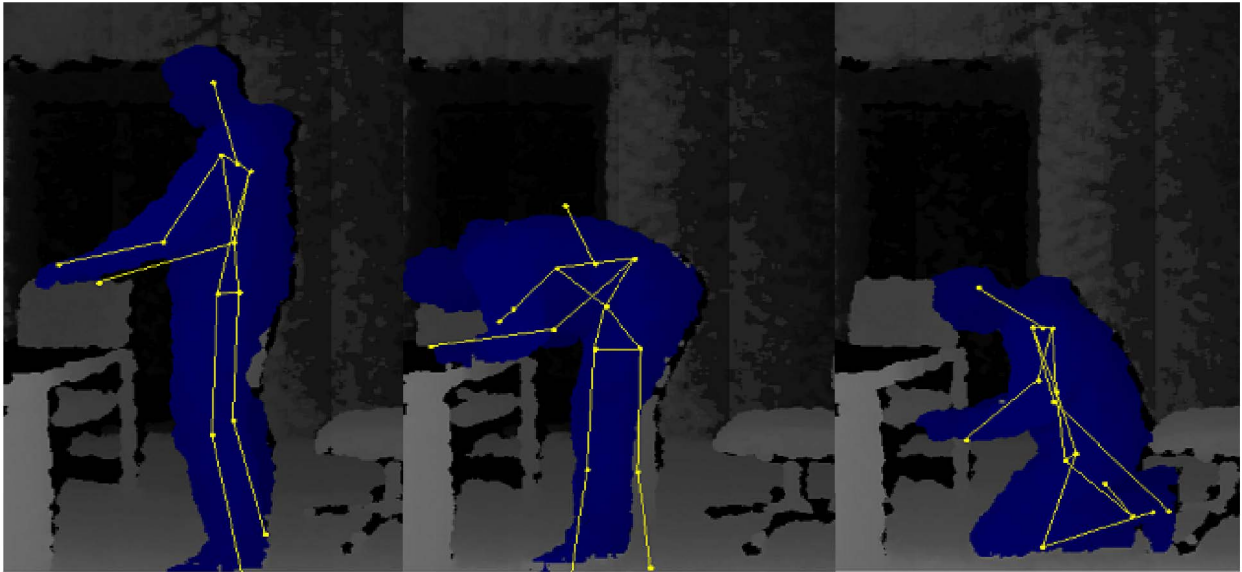


Figure 5.6: Shortcomings of the joint localisation algorithm of Shotton et al. [11] and their implementation in the Primesense NiTE library. The middle picture suggests that the posture adopted was not present in the training data set and, thus, cannot be recognised.

incorrect. Figure 5.6 provides examples where the algorithm in Shotton et al. was applied on postures and viewpoints that differed from the training set.

Filtering Approaches

Filtering or tracking approaches [138] assume a temporal coherence between successive results and work best when the motion between image frames is small. Gall et al. [131] described the filtering task as the estimation of an unknown state, the model parameters, from a sequence of noisy observations. One can imagine that the algorithms filter out the noise from the signal I such that only the state information \mathbf{q} remains. The core idea of filtering algorithms is to assume that the state sequence \mathbf{q} follows an underlying stochastic process.

The techniques to solve this filtering problem can be divided into single hypothesis and multiple hypothesis methods. Single hypothesis methods [139] are based on the Kalman filter [140] or the extended Kalman filter and output only the most probable parameter set for the current frame. The model parameter estimations of the upcoming frame are strictly based on this hypothesis. The benefit of single hypothesis approaches is that there is a closed-form solution for the underlying mathematical problem, which can be efficiently determined. The main disadvantage is that incorrect or inaccurate estimations are propagated to further frames. In other words, once the system goes into an erroneous state, it cannot recover. Multiple hypothesis approaches [141, 142, 143] are based on sequential Monte Carlo methods. As opposed to single hypothesis methods, they propagate multiple guesses over time. Each hypothesis is assigned a weight that represents its probability of being the correct estimation. For the current estimation, the hypothesis with the heaviest weight or the mean of all hypotheses is computed. Multiple hypothesis-based

approaches are capable of better recovering from failures than single hypothesis methods because the “backup hypotheses” are kept, which can be used when the current one has shown to be wrong. Moreover, the weighting makes certain that incorrect hypotheses have little influence on the final estimation. Yet, multiple hypothesis methods are more computationally demanding than single hypothesis algorithms.

The key advantage of filtering methods is that they use additional information from previous frames. This can make the estimation of the parameters more robust. However, single hypothesis methods are especially susceptible to incorrect estimations in the past frames. Moreover, if the initial frame contains a non-optimal estimation, all subsequent frames will base their estimations on the incorrect frame. Multi-hypothesis approaches attempt to provide more robustness, however, with growing parameter space, a large amount of hypotheses have to be generated and evaluated.

Generative Approaches

Generative approaches operate following the analysis-by-synthesis principle[12]. The objective is to find the model parameters that best explain the image data. Generative approaches require two components to be designed. First, an operation has to be developed that creates a model image out of the current model parameters. Moreover, an objective function has to be designed representing the similarity or dissimilarity between the input image and the model image. The generative method strives to determine model parameters that optimise the objective function. In this thesis, the convention is maintained that an objective function shall be minimised. For the optimisation process, local search strategies or global search strategies are applied.

Local search methods [144, 145, 146, 147, 148] are based on computing the derivatives of the objective function subject to the parameters and iteratively updating the parameters towards directions where the objective function decreases. Local methods achieve solutions which are very near to a local optimum. However, the main downside is that the solution can be a local minimum of the objective function. A local minimum can still be a “bad” solution as it leads to an unacceptably high objective function value. In order to reach a minimum that yields an objective function value similar to the global minimum, a parameter initialisation near the target is essential. For promising initialisations, generative approaches are combined with discriminative or filtering frameworks. The resulting methods are hybrid approaches. As an example, initialisations can be realised from the estimation of the previous frames [146] (similar to filtering). Discriminative and generative approaches were combined in Sharp et al. [132], where the authors used the algorithm proposed by Shotton et al. [11] to create initial hypotheses and refined them with a generative method. Alternatively, specialised part detection algorithms can be applied. Qian et al. [149] detected finger tips and Elhayek et al. [150] localised specific body parts to guide subsequent generative algorithms.

Global search strategies attempt to overcome the problem of local minima by discretising the search space [151] or applying multiple hypothesis optimisation methods

[152, 153, 154, 155, 156]. Discrete methods ensure the global minimum of the search space is limited to a finite set of candidates. However, depending on the sampling rate of the search space, the accuracy might not be sufficient for a given task or may only be achievable through exhaustive computational effort. What global optimisation strategies have in common is that they are computationally intensive. Similar to multiple hypotheses filtering approaches, the number of objective function evaluations is many times higher than in local search approaches. Only recent parallel computing architectures have made it possible to achieve real-time performance with global search strategies.

Generative approaches can deal with fast movements or can estimate postures from a single image when a reliable initialisation is available. Moreover, they can address the variance of motions through automated search of the parameter space. This makes tedious manual generation of training examples, like in the case of discriminative approaches, obsolete. On the other hand, the computational cost of generative approaches is high because of the frequent creation of the model images. Moreover, the quality of the estimations is highly dependent on the initialisation of the hypotheses.

Selected Approach

As described previously, discriminative approaches require a large amount of labelled data. Obtaining such data is too expensive, which is why this approach was discarded. Filtering and generative approaches seemed to be comparably attractive. However, looking at recent publications related to joint localisation from 2011 - 2016, generative [132, 146, 148, 152, 154] and discriminative [11, 132, 137, 135, 136] approaches were clearly preferable to filtering [141]. This is why the generative approach was elected for the joint localisation component. For the developed 2D approach, a discrete method was used, which does not require any initialisation. In the 3D joint localisation approach, generative localisation was combined with an initialisation through previous estimations (filtering-like approach). Table 5.2 summarises the criteria used for the choice of the joint localisation approach.

Table 5.2: Overview of joint localisation alternatives. Columns highlighted in green represent the selected approach.

Criterion	Discriminative	Filtering	Generative
Training data required	yes	no	no
Computational complexity	low	low - high	high
Dependent on initialisation	no	beginning of recording	yes (except discrete methods)

Risk Analysis

The problem of localising human joints in an image is still not fully solved. In the current state of the art, there is no general method, which provides robust estimations in all situations. Rather, acceptable quality of the joint localisations has been reported for special cases with controlled setup. Whether the posture assessment task can be transformed into one of these fortunate cases remains a subject of investigation. Since the objective functions used in generative methods are normally too complex to derive or guarantee mathematical properties, the only way to evaluate the feasibility of this approach is conducting experiments. Not only is achieving sufficient accuracy and robustness a challenge, but maintaining the runtime of generative approaches at a few seconds per frame requires the application of elaborate algorithmic transformations and parallelisation techniques.

5.2.3 Image Acquisition Devices

This section discusses which image acquisition devices could be used for generating the input data for the localisation algorithm. The first question answered is whether to use colour (RGB) or depth images (see fig. 5.7). RGB images contain 2D information of the scene. Depth images are considered as 2.5D information, as a single image cannot capture the subject from all perspectives. However, the captured part contains 3D data. In order to obtain "real" 3D data, multiple RGB or depth images from different viewing points have to be merged. In the scope of this thesis, only algorithms operating on a single image were considered. Poppe [128] stated that solving the joint localisation task on 2D data leads to ambiguities, which can be hard to algorithmically resolve. Hence, the depth data format was preferred as it better preserves the 3D information than RGB images. Another problem was that RGB images have more variance in situations of different environmental illumination or clothing than depth images unnecessarily increasing the difficulty of the localisation task. Finally, recent successful joint localisation methods mostly count on 2.5D depth data [11, 132, 137, 135, 136, 148, 152, 132]. Bearing in mind the arguments given, the search was concentrated on modalities that generated depth images.

There are a variety of approaches to acquiring depth images. The most common techniques are discussed with regards to their suitability for the given joint localisation task. This list is not claimed to be complete. Among the so-called shape-from-x methods, where x denotes the features used,

- Time-of-Flight (ToF)
- Shape from focus
- Shape from shading
- Shape from texture
- Photometric stereo



Figure 5.7: Colour image (left) and corresponding depth image (right) from the Microsoft Kinect® sensor.

- Structured light
- Stereo vision
- Shape from motion
- Light field technology

were discussed.

Pre-Selection of Relevant 2.5D Image Acquisition Modalities

In order to find candidate devices, the method of elimination was employed. In brief, options were eliminated for three reasons: high cost, low frame rate and infeasible environmental conditions: Initially, the light field cameras were discarded because acquiring the hardware was financially rather expensive. As an example, the company "Raytrix" offered in 2012 light field cameras starting from around 4000 Euros without lenses. Further, there were methods abandoned because the achievable frame rate was too low - depth from focus is an interesting approach, but requires a long period of time for the generation of the focal stack, the succession of images with different foci utilised to determine the depth information. In addition, the computation itself is remarkably lengthy. Suwajanakorn et al. [157] observed an overall computing time of around 20 minutes per image on a single CPU in a recent publication. Finally, there were methods needing exclusion because the given use case could not guarantee the environmental conditions they required. Photometric stereo approaches need multiple images with different lighting. Tracking 3D motion would require quickly altering, through different lighting, setups in the scene. This continuous switch would rapidly confuse the worker. Shape from shading is mostly used in controlled environments where the lighting parameters are known [158]. Even recent approaches, such as Han et al.'s [158], require objects with uniform albedo. Unfortunately, determining the lighting or guaranteeing the colour of the object of interest to be uniform is mostly not possible for real world images. Shape from

texture requires a regular texture on the scene, which can also not be definite at an assembly workplace. Last but not least, the use case contained many situations where the subject was static. As there was only a little motion in the scene, the depth from motion approach was expected to encounter problems. Having eliminated most of the methods, there were three viable options to obtain depth information from the scene: stereo vision, structured light and ToF. In the following, the principles and their characteristics will be briefly described.

Stereo Vision

The stereo vision approach obtains depth information by reconstructing a 3D point from the corresponding 2D points from two different views. Stereo cameras consist of two mono or RGB cameras often placed in parallel. The depth of a point in the scene can be reconstructed when it is visible by both cameras. The reconstruction process is called triangulation. The most important step for stereo cameras is to find the points in both images that correspond to the same 3D point. This process is called stereo matching. Scharstein and Szeliski [159] published the Middlebury dataset for comparing stereo matching algorithms with one another. On the webpage www.middlebury.edu/stereo mentioned in their paper, a continuously updated overview of stereo matching the techniques can be found. In general, the characteristics of the stereo vision systems, e.g., Field of View (FOV) of the depth image or depth resolution, depend on the characteristics of the cameras used. The depth resolution can be computed using the following formula [160]:

$$dZ = \frac{dD \cdot Z^2}{f \cdot B} \quad (5.1)$$

As can be seen in Eq. 5.1, the size of a depth element, dZ , which corresponds to the depth resolution, is dependent on the distance of the 3D point to the cameras, Z , the focal length of the cameras, f , and the baseline between the cameras, B . Moreover, the achievable disparity resolution, dD , is an essential factor. This resolution depends on the stereo-matching algorithm employed and can either lie in pixel or sub-pixel range. The biggest advantage of stereo cameras is that they are a passive technology. Hence, multiple stereo cameras from different viewing angles do not interfere with their respective signals, unlike that which occurs with active technologies. Additionally, high-depth resolutions can be achieved, even at greater distances, if the resolution of the cameras, the baseline or the focal length are high. The main disadvantages of stereo cameras are that varying the illumination or a lack of texture in the scene can disturb the reconstruction. Moreover, a time-consuming calibration process is often needed. Last but not least, the computational complexity of stereo-matching algorithms is high, which adds to the time to the overall computation.

Structured Light

Structured light systems follow an active stereo approach. The underlying idea is to project a known pattern onto the scene that becomes distorted by the scene geometry.

One notable example of a structured light sensor is the Kinect. Since its release, the Kinect has dramatically advanced the area of computer vision [161]. The reasons for this trend are low price, not needing calibration of the system and low computational costs in the depth-generation process. Moreover, on the software site, Shotton et al. [11] accomplished a breakthrough in the field of real-time motion analysis with this device. Yet, the Kinect has its shortcomings in terms of accuracy. Especially for distances above $4m$, their area of application is limited because of the low-depth resolution and, additionally, the high variance in depth values. According to Khoselsham and Elberink [162], the depth resolution at distances of $4m$ is $5cm$ with a standard deviation of $2.5cm$. Furthermore, this technology does not allow for the creation of a dense depth image with a measured value for each pixel. According to Sarbolandi et al. [163], the depth information for each pixel is estimated from the local image information in a 7×7 or 9×9 window. Finally, when using multiple Kinects to resolve occluded parts of the body, one has to contend with the problem that the sensors interfere with each other. The interference of signals results in dead spots within the depth images. Resolving this problem requires modifying the hardware, such as multiplexing of the sensors [164] or periodically shaking the devices [165, 166].

Time-of-Flight

Similar to the structured light approach, the ToF technology is also an active technique. The main rationale is to emit light onto the scene and measure the time it travels to the object and back to the sensor array. The travel time can then be converted into a depth value [163]. After the release of the first Kinect sensor, Microsoft refined its depth sensor technology and altered the depth image acquisition principle from structured light to ToF. Consequently, the Kinect v2 or Kinect One was released in 2014.

After its release, the Kinect v2 was investigated by several groups [167, 163]. The Kinect v2 offers a higher FOV (70.6° in the horizontal direction and 60° in the vertical direction) than its predecessor. Furthermore, Yang et al. [167] observed the depth resolution to be higher, especially at longer distances ($4mm - 5mm$ at a distance of $4m$). With this all being said, the Kinect v2 has the same advantages as its predecessor. Additionally, it corrects the weakness of low accuracy. The device now creates a dense depth map that permits a detailed analysis of the scene. Nevertheless, vulnerabilities in terms of particular environmental conditions continue [163].

Selected Approach

Among the alternatives, the first generation Kinect was chosen for depth image acquisition. This modality was preferred to stereo vision approaches because no calibration is required and the computational complexity for depth image acquisition is significantly lower, leaving more capacity for the motion analysis algorithms. The Kinect v2 is an even more interesting alternative than the first Kinect, however, at the time of the algorithm design, it had not yet been released. Nevertheless, because the methods developed are

device independent, the shift to the Kinect v2 could be easily realised. Table 5.3 summarises the criteria for the selection of the image acquisition device.

Table 5.3: Overview of image acquisition device alternatives. Columns highlighted in green represent the selected approach.

Criterion	Stereo	Time-of-Flight (Kinect v2)	Structured Light (Kinect v1)
Calibration required	yes	no	no
Computational complexity	high	low	low
Availability at development time	yes	no	yes

Risk Analysis

As researchers [11, 148, 152] have shown, the accuracy of the Kinect v1 is sufficient to localise joints in distances below 3m. However, the use cases discussed in the paper enabled the Kinect to be placed at a rather small distance in highly controlled environment. It has to be examined whether the technical properties of the Kinect v1, such as resolution and robustness to external influences, satisfy the application case of the present work.

5.3 Posture Classification Workflow

Section 5.2 has derived how the algorithmic base for the posture assessment method would look. First, the section has argued that an indirect posture classification method was most suitable for the given task because the upcoming posture optimisation method required 3D joint coordinates, anyway. Using indirect features, a joint localisation method had to be selected, which obtains these coordinates. The generative approach was selected as it does not require a massive amount of training data and has proven to be effective according to recent literature. Finally, the question which image acquisition device to use has been discussed. The selected device was the first generation Kinect sensor as it was cost-efficient, simple to use and does not require a massive amount of computing resources.

In the upcoming sections, the concept of the posture classification algorithm is presented. The workflow is depicted in Figure 5.8. The method contains three parts: camera setup, joint localisation and posture classification including EAWS computation.

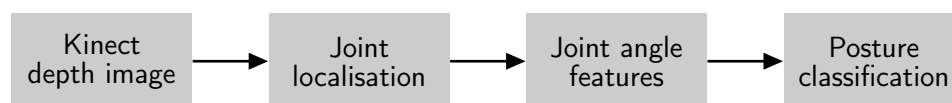


Figure 5.8: Illustration of the developed posture classification workflow.

5.4 Camera Setup

At the beginning of the pipeline is the depth image. This section tackles the question of which viewpoint to select. The answer was derived in three steps: 1) the theoretically required minimal spatial resolution (in mm or cm) of a 2.5D camera to properly recognise each posture was determined; 2) the spatial resolutions possible with a camera setup featuring the Kinect was discussed; and 3) using required and achievable spatial resolutions, possible options for the placement were evaluated. For the explanations throughout the sections, the camera coordinate system notation depicted in Figure 5.9 is used.



Figure 5.9: Camera coordinate system notation used in this thesis.

Required Spatial Resolution

First, the theoretical spatial resolution of the data required to properly use the posture assessment tools, was calculated. For the analysis, the exemplary ergonomics assessment tools, REBA [29] and EAWS [34], were chosen. With these tools, the posture classes are identified from the values of the joint angles. The starting point of the analysis was to determine the smallest joint angle resolution required in order to be able to properly distinguish each posture class. Both tools work in such a way that a posture class is determined by identifying the combination of joint angle intervals. The assessment tools define fixed intervals for each joint angle. An exemplary visualisation of joint angle intervals for the back joint can be seen in Figure 5.10a. In EAWS, a back angle in the green area ($< 20^\circ$) means an upright posture, in the yellow area ($20^\circ - 60^\circ$), a bent-forward posture, and in the red area ($> 60^\circ$), a strongly bent-forward posture. Additionally, the angles of the lower limbs define whether the subject is sitting or standing bent forward. Finally, the angles of the shoulder joint decide whether the arms are below, at or above shoulder level. Thus, a posture is basically a combination of joint angle intervals.

This posture class representation makes it obsolete to confirm the spatial resolution for each posture in isolation. Rather, it suffices to consider the intervals for each relevant joint. Table A.1 in the Appendix lists the angle intervals and minimal angular resolution required (α_{min}) for EAWS and REBA. As can be observed, the EAWS is less precise than REBA in the description of the angular classes. Taking into account the manual usage of such a worksheet where the observing practitioner has no opportunity to determine the exact angular values, this representation seems to be more intuitive. For automating this

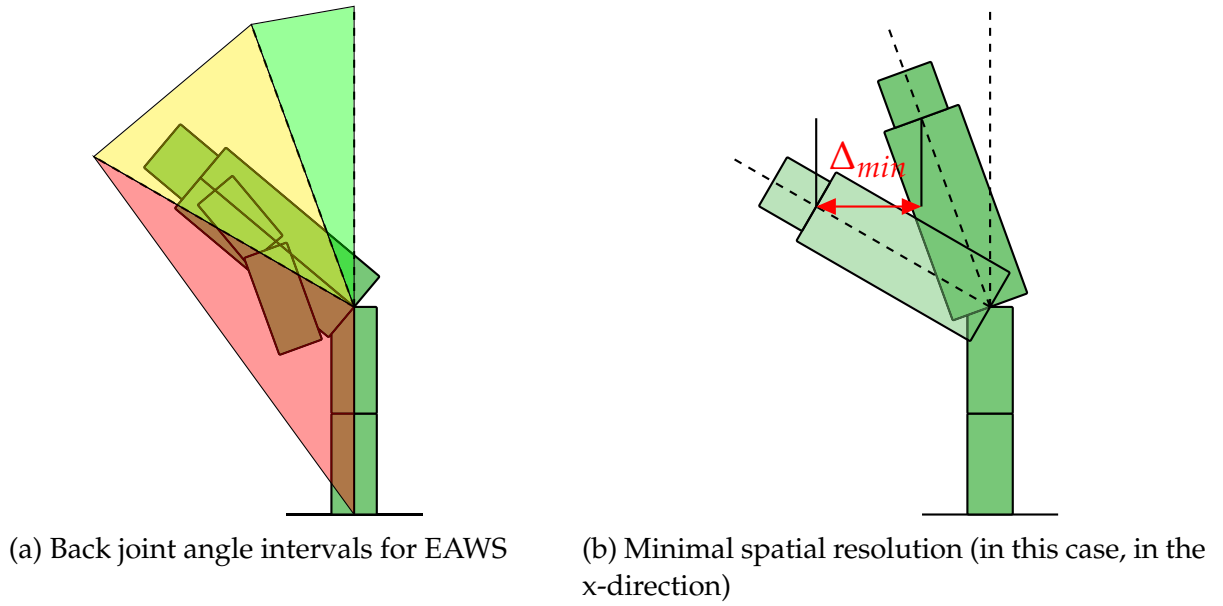


Figure 5.10: Visualisation of variables influencing the required spatial resolution.

assessment, however, the ill-definition of angles posed a challenge. In order to obtain a required spatial resolution from the minimal angular resolutions, the lengths of the body segments had to be considered. The segment lengths were derived from the German standard, "DIN 33402-2:2005-12 - Ergonomie - Körpermaße des Menschen - Teil 2: Werte" [168], which posits typical anthropometric values for the 18 to 65-year-old population in Germany. The length values of particular segments for the 5th percentile (smallest 5%) are depicted in Table B.1 in the Appendix. Using the minimal angular values, α_{min} , and the lengths of the body parts, l , the minimal required spatial resolution, Δ_{min} (see Fig. 5.10b for visualisation) can be calculated as

$$\Delta_{min} = l \cdot \sin \alpha_{min} \quad (5.2)$$

Table A.1 in the Appendix shows the calculated Δ_{min} for each part of the body with each tool.

Achievable Spatial Resolution with a Kinect

In the next step, the spatial resolutions which could be theoretically captured with a Kinect were determined. The properties of the Kinect have been intensively investigated by researchers [162, 163]. The depth resolution lies in the magnitude of millimetres when the distance to the scene is below 1m. However, it quickly decreases with increasing object distance. The same applies for the accuracy of the depth values, which drop to 4cm at a distance of 4.5m. The depth resolution of the image acquisition device was significantly lower versus the resolution in the x or y directions (horizontal or vertical). Hence, the depth component was focused on. The depth resolution of the Kinect depends on its distance to the scene. The resolution decreases the further the Kinect is placed. In order to decide whether the Kinect is able to attain the required depth resolution, the minimum

distance of the Kinect to the scene where it is still possible to fully capture the worker, was calculated. The dimensions of the minimum plane to be captured were defined with -2m to 2m in the x-direction and -1.25m to 1.25m in the y-direction. Given the minimum extent to be captured, a , and the FOV of the device, α_{FOV} , the minimum distance, z_{min} , could be computed as follows (see Figure 5.11 for visualisation):

$$z_{min} = \frac{a}{2 \tan \frac{\alpha_{FOV}}{2}} \quad (5.3)$$

In the present case, the minimum extent was $a_x = 4m$ in the x-direction and $a_y = 2.5m$ in the y-direction. The Kinect has a FOV of 58.5° in the x-direction and 46.6° in the y-direction. Inputting these values into Eq. 5.3, the Kinect had to be positioned approximately $z_{min} \approx 3.5m$ from the scene.

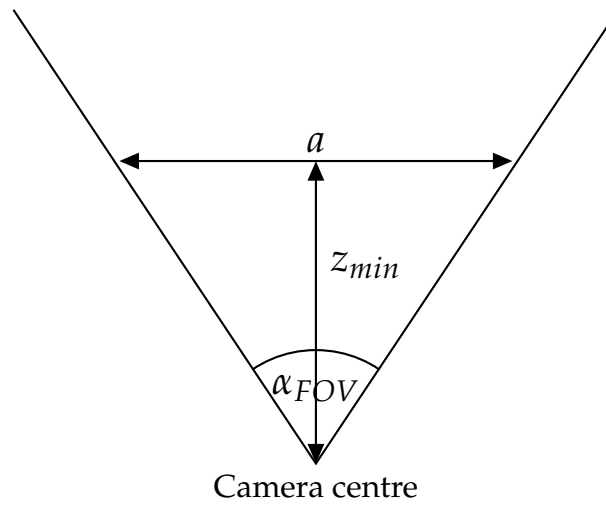


Figure 5.11: Illustration of the variables z_{min} , α_{FOV} and a .

At this distance, z_{min} , the spatial resolution (smaller than 5cm) including variances of 2cm [162] theoretically were sufficient for capturing the features for almost all postures in the EAWS. However, postures relating to lateral bending or rotation of the trunk would have been hard to distinguish when only considering depth data. Within the scope of this thesis, detecting these postures is not covered. This leads to a potential loss of 30 points in the EAWS and four points in RULA and REBA.

Evaluation of Viewpoint Options

Apart from the choice of the distance to the scene, the viewpoint of the camera was another design alternative necessitating consideration. There are four options for choosing the observation perspective: observing from the front or back, the side, the top or an isometric perspective, where the latter is a mixture of the aforementioned options. The front perspective has the disadvantage that bigger workpieces to be processed by the worker will occlude the view of the human. From the back perspective, the trunk itself will often obstruct the view on the upper limbs. The top perspective will create a situation such

that the most distinctive features of the body, such as orientation of the back, will heavily rely on the depth resolution. Finally, the isometric perspective would be a fair compromise, however, 3D joint localisation algorithms are definitely required in this case, which are more challenging to develop than 2D joint localisation methods. As lateral bending and trunk rotation were not considered, the side perspective was selected - it captures the most distinct features of a posture with the highest possible resolution along the x-axis.

5.5 Joint Localisation

Having obtained the depth image from the side perspective, the next step was to apply the joint localisation algorithm. Although there are existing implementations of joint localisation algorithms for the Kinect [169], experience gathered from pre-experiments suggested that they fail to provide robust and accurate joint position estimations during application (see Fig. 5.6 for examples). Thus, an own joint localisation algorithm was developed tailored to the given requirements. This section describes the methods developed to determine the 3D positions of the joints from a given depth image. First, the pre-processing steps to support and simplify the localisation problem are explained. Afterwards, the mathematical notation used throughout the algorithmic descriptions is introduced. Finally, the approaches themselves are described. Two methods have been developed which both assume that there is only one subject in the observed space.

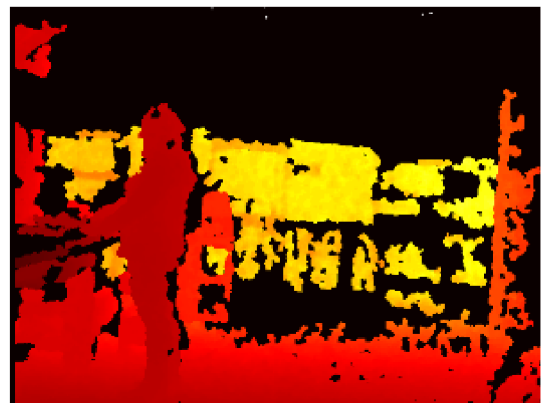
A 2D joint localisation approach was selected. The rationale behind was that for determining most postures of the EAWS, 2D information is sufficient. Besides, in comparison to generative 3D joint localisation approaches, the number of parameters to be determined is lower, which simplifies the core optimisation problem. The 2D joint localisation method assumes that all motions of the worker are performed along the x-y plane. This assumption does not impede the task of recognising the 2D postures of the EAWS. However, it does reduce the model parameters for determination. As a consequence, the algorithm becomes more robust and computation is more efficient. Furthermore, a 3D joint localisation method was developed based on recent approaches using the Particle Swarm Optimisation (PSO) method [152]. Designing this method followed the intention to test a state of the art baseline according to its suitability for the given posture assessment task. Both methods follow the generative localisation approach. This framework contains four components [12]: image features, model, model image and distance measure between image features and model image. The realisation of these components will be separately described for each localisation method. In addition, the method used to optimise the model parameters to minimise the distance measure is explained.

5.5.1 Pre-processing of the Image Data

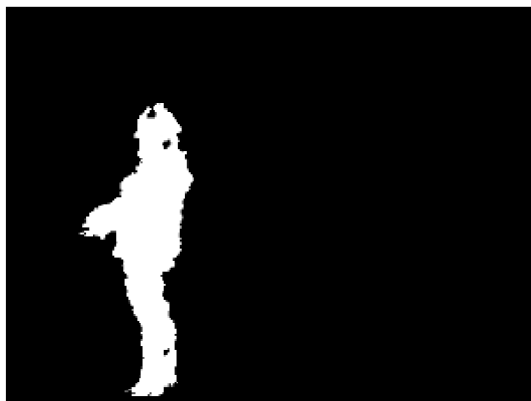
As input data, the localisation algorithm provides a depth image acquired by the Kinect (see Figure 5.12b). However, the raw image does not suit the task as it contains a high amount of information about the background scene. Therefore, the first task was to segment the human out of this scene. Fortunately, this human segmentation image, which is referred to as a user mask, could be robustly obtained by background subtraction applied on the depth image (see Figure 5.12c). The built-in routine from the NiTE2 middleware [169] was used, which was shipped with the OpenNI2 framework [170]). This implementation is able to provide segmentations of up to six subjects. The only requirement is that the subject has to make an initial motion before the segmentation starts. It suffices that the subject enters the scene or simply moves forward. Having obtained the binary user



(a) Colour image



(b) Depth image



(c) Masked image



(d) Image to determine the foot point

Figure 5.12: Intermediate results of the pre-processing pipeline.

mask and the masked depth image, the depth image was thresholded in the x , y and z directions. The reason for this was to remove subjects observed by the sensor but located outside of the space to be monitored. As such, the system could focus on the worker in

the monitored space. In order to obtain the 2D point where the foot was located, morphological erosions followed by dilations were first performed on the mask to remove speckles and decrease noise on the mask (see Figure 5.12d). After that, the image was scanned from the bottom looking for the first row where there is a foreground pixel. A connected component analysis was run on all pixels located in this row. The centroids of the components obtained represented the foot-point candidates. In most cases, there was only one component. In case where there were multiple components, the leftmost one in the image was taken. The pre-processing pipeline is summarised in Figure 5.13.



Figure 5.13: Illustration of the pre-processing workflow.

5.5.2 Notation

In the following, the skeleton to be obtained is represented by a kinematic model symbolised by \mathcal{M} . The parameters of \mathcal{M} , namely the rotational angles of the joints, are composed into the parameter vector \mathbf{q} . Moreover, each segment of the skeleton is covered by a "wrapper" \mathcal{S} , representing an approximation of the subject's shape. As \mathcal{S} can be an arbitrary shape, their parametrisation varies. The model and wrapper are depicted in Figure 5.14.

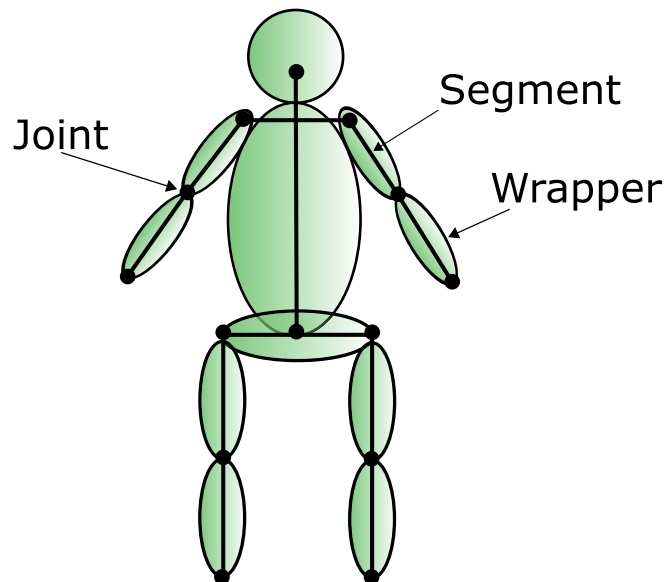


Figure 5.14: Human model terminology used in the following sections.

Referring to the input data, the notation I will be used to indicate a 2D image. I can either be a depth image, a binary mask (where 0 denotes a background and 1 represents

a foreground pixel) or an RGB image. If indexing is relevant, the notation $I(x,y)$ will be employed. Moreover, the operation, $|I(x,y)|$, denotes the number of non-zero pixels of I . For the sections, 5.5.3 and 5.5.4, operations will be specified that create a model image $I_{\mathcal{M}}$ from a given model \mathcal{M} . Finally, a distance measure between two images $d(I_1, I_2)$ is defined, which returns 0 if $I_1 = I_2$ and bigger values the more I_1 differs from I_2 . This distance measure is then minimised yielding the optimal parameters \mathbf{q}^* .

5.5.3 2D Joint Localisation

The 2D joint localisation algorithm acquires the silhouette image, which will be called user mask, as input and returns joint angles along the camera x-y plane for foot, knee, hip, shoulder and elbow (see Fig. 5.15). This approach was inspired by the method of Li et al. [151].

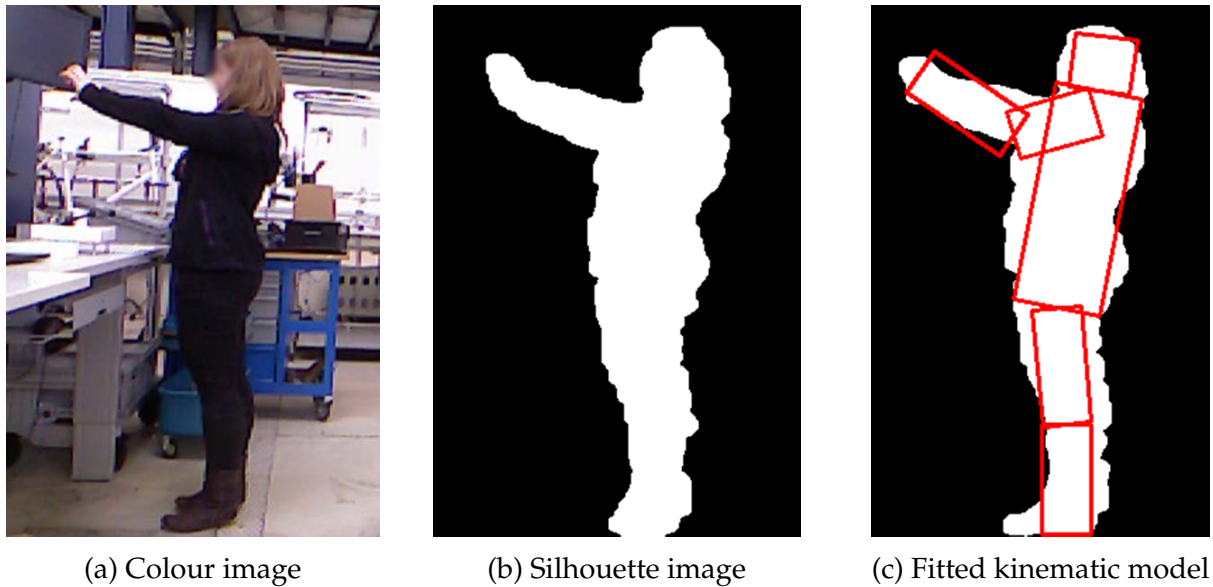


Figure 5.15: Workflow of the 2D joint localisation approach.

Image Features

As image features, the pre-processed user mask containing the worker pixel region (see Fig. 5.15b), was used. The generation of the user mask was described in Sec. 5.5.1. The core assumption of this approach is that the 2D postures can be purely recognised by 2D silhouette information. Adding 2.5D information might have helped, but because of the poor accuracy of the Kinect sensor, there was the danger that more noise would be added than usable information.

Model

A human model was composed from two kinematic chains (see Fig. 5.16a):

- $\mathcal{M}_{Body} = (\mathcal{S}_1, \dots, \mathcal{S}_4)$ and
- $\mathcal{M}_{Arm} = (\mathcal{S}_5, \mathcal{S}_6)$.

\mathcal{M}_{Body} (depicted in green) represents one lower limb, torso and head. \mathcal{M}_{Arm} (depicted in red) indicates one upper limb. Only the more highly located upper limb and one lower limb can be tracked. Though simplified, this model sufficed to provide enough information for ergonomic assessments using EAWS. In the EAWS, only the more highly located hand and its corresponding upper limb are of interest. These parts represent the more significant physical load. The joint localisation procedure is able to capture this limb no matter whether it is the arm facing the camera or the other one. As soon as the part of the body is visible in the silhouette, it will be localised. Even if both arms are visible, the optimisation routine can be programmed to prefer solutions where the higher arm is captured.

For lower limbs, there are two commonly occurring cases - legs in identical poses (sitting, kneeling, standing) or one leg stepped forward (standing). In both scenarios, the physical load can be determined analysing only one of the two legs. As can be seen later, the proposed algorithm mostly certifies that only the limb with the largest load is tracked. For the 2D model, the angular parametrisation depicted in Fig. 5.16b was employed.

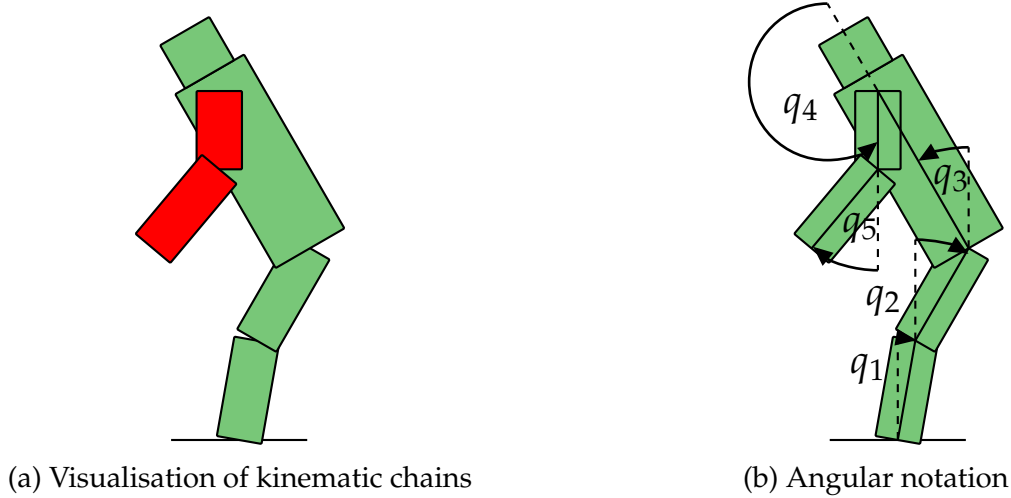


Figure 5.16: Visualisation of the 2D human model.

Model Image

The model image was created by generating the silhouette of the kinematic chain. The wrappers were modelled as rectangular boxes. The parameters, w_j and h_j , of wrapper, \mathcal{S}_j , denote its 2D width and height. The angle, \mathbf{q}_j , signifies the rectangle's 2D orientation:

$$\mathcal{S}_j = (\mathbf{q}_j, w_j, h_j) \quad j \in 1..J \quad (5.4)$$

A silhouette image I_M can be computed by inspecting whether each pixel lies inside of one of the two triangular segments of the rectangle (see Figure 5.17). This inclusion check is accomplished by transforming the 2D pixel location into barycentric coordinates [171] of the corresponding triangle and determining whether every barycentric coordinate is found in the interval between 0 and 1. In order to further speed up the process, the set of pixels to be checked can be reduced to the pixels lying in the region of the maximum and minimum 2D coordinates of the rectangle.

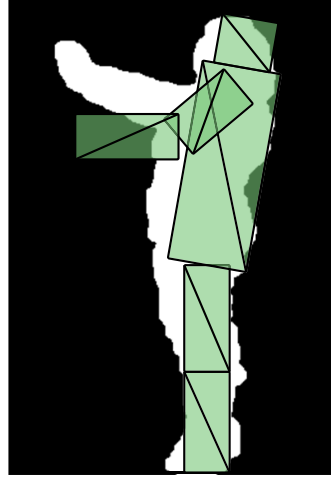


Figure 5.17: Silhouette image and overlaid model triangles.

Distance Measure

The distance measure considers the overlapping region between the silhouette and model image. The region can be expressed by the ratio between two pixel set sizes:

$$d_{2D}(I, I_M(\mathbf{q})) = 1 - \frac{|I \cap I_M(\mathbf{q})|}{|I \cup I_M(\mathbf{q})|} \quad (5.5)$$

The term, $|I \cap I_M(\mathbf{q})|$, contains the number of pixels that have the value 1 in the silhouette image, I , as well as in the model silhouette I_M . Therefore, it represents the intersection of both silhouettes. The term, $|I \cup I_M(\mathbf{q})|$, constitutes the size of the union of both pixel sets. To garner a distance measure where low values represent highly similar images, the computed ratio was subtracted from the value 1.

Optimisation

The localisation task can be defined as finding the optimal model parameters \mathbf{q}^* to the optimisation problem.

$$\mathbf{q}^* = \arg \min_{\mathbf{q}} d_{2D}(I, I_M(\mathbf{q})) \quad (5.6)$$

A discrete optimisation strategy was applied to obtain the optimal solution. In comparison to continuous optimisation procedures they might achieve lower accuracies. An example case where the discrete sampling of candidate solutions impairs accuracy can

be observed in Fig. 5.15c. The upper arm is slightly bent although the subject in the image is reaching forward. Due to the sampling of the candidate space, the algorithm has to decide between suboptimal candidates. Nevertheless, a big disadvantage of continuous optimisation approaches is that they cannot make certain that the solution is a global minimum of d_{2D} . Contrastingly, discrete optimisation methods can deliver a globally optimal solution for a finite solution set. In the case of posture classification, a high accuracy advantage is not necessary as a minimal angular resolution is provided (see Tab. A.1 Appendix). The drawback of discrete optimisation approaches is that they are computationally expensive. The approach requires drawing candidate samples from the continuous solution space of each joint angle. A naïve solution algorithm would have to evaluate all combinations of joint samples to find which is optimal. The size of the candidate solutions grows exponentially with respect to number of parameters. If the number of parameters is n and the number of samples per joint is m , the number of combinations to evaluate is m^n . Although computationally demanding, this approach was still preferred for two reasons - the number of parameters, n , was only 5 for the given model. Besides, the problem could be modified to attain significant speed enhancements.

The algorithm was sped up by approximating the original distance measure, d_{2D} by a distance function $d_{Approx2D}$, which corresponded to the mean overlap per wrapper. First, the width, w_j , and the height, h_j , of a wrapper, \mathcal{S}_j , were assumed to be known. The modified distance function compared the input silhouette I with each model image of an isolated segment, $I_{\mathcal{S}_j}(\mathbf{q})$, and computes the mean of the obtained distances:

$$d_{2D}(I, I_{\mathcal{M}}(\mathbf{q})) \approx d_{Approx2D}(I, I_{\mathcal{M}}(\mathbf{q})) = \frac{1}{J} \sum_{j=1}^J \left(1 - \frac{|I \cap I_{\mathcal{S}_j}(\mathbf{q})|}{w_j \cdot h_j} \right) \quad (5.7)$$

The transformed optimisation problem:

$$\mathbf{q}^* = \arg \min_{\mathbf{q}} d_{Approx2D}(I, I_{\mathcal{M}}(\mathbf{q})) \quad (5.8)$$

is comprised of a sum of single terms where each only depends on the previous model parameters of the kinematic chain. For instance, the pose of a back segment (\mathcal{S}_3) can be established by only considering the foot (\mathbf{q}_1), knee (\mathbf{q}_2) and back (\mathbf{q}_3) joint angles. The joint angle parameters for shoulder and elbow are not required. For further illustration, refer to Fig. 5.16b. In general, the pose of \mathcal{S}_j can be computed by just considering \mathbf{q}_k with $k \in 1..j$. This functional property was exploited by applying an improved optimisation strategy. Instead of evaluating all combinations of parameters, all alternatives for the foot angle were assessed first and the results saved. Next, the saved values were reused for all alternatives containing the foot and knee. The next joint was then subsequently added in the kinematic chain. The analysis for the upper limb model was performed similarly. This scheme led to a lower number of evaluations. Taking advantage of this alternative distance function, the optimisation problem was transformed into a shortest path problem in a graph, which can be efficiently solved by dynamic programming algorithms. The

n -th joint angle sample of joint j is denoted with:

$$q_{j,n} \quad j \in 1..J, n \in 1..N \quad (5.9)$$

Out of all angle values in the solution set, $q_{j,n}$, a graph is constructed: Let $G = (V, E)$ be this graph represented by a set of vertices, V , and a set of edges, E . The vertices and edges are constructed as follows:

- Each $q_{j,n}$ is represented by a vertex, $v_{j,n} \in V$. Hence, the set of vertices represents all possible angle values for each joint. For $j > 1$, there are multiple copies of $q_{j,n}$. The number of copies is N^j (see Fig. 5.18 third node row for example). The graph represents a search tree containing all possible combinations of joint angles.
- There are edges connecting vertices that have neighbouring joint IDs. An edge connects $v_{j,n}$ with one copy of $v_{j+1,n} \forall n$.
- The distance between the silhouette image, I , and the model image of a single wrapper, I_{S_j} , with the parameters, $q_{1,n}, \dots, q_{j,n}$, is defined as:

$$1 - \frac{|I \cap I_{S_j}(q_{1,n}, \dots, q_{j,n})|}{w_j \cdot h_j}$$

This value is assigned to the edge weight pointing to the corresponding vertex (see Fig. 5.18 for an exemplary edge and its weight coloured in red).

- Finally, a start vertex $v_s \in E$ and a terminal vertex $v_t \in E$ are added to the search tree. The start vertex has no edges pointing to it, but all edges from v_s point to the vertices representing the first segment of the kinematic chain ($v_s, v_{1,n}$). The same scheme applies to the terminal vertex. Here, G has only edges ($v_{J,n}, v_t$).

A path in G from v_s to v_t symbolises a legal model configuration, where its length (sum of all edge weights on the path) represents the overall cost $d_{Approx2D}$. Finding the shortest path or the configuration with the lowest cost can be realised by the A* algorithm [172]. The advantage of A* search is that it does not need the whole graph in advance. Rather, the nodes and edges are iteratively built when they appear to be relevant to figuring out the shortest path. Hence, only a fraction of nodes and edges are generated to find the minimum value. This speeds up computation and reduces memory consumption.

A set of extensions was applied to the algorithm that demonstrably improved accuracy, robustness and computation time of the localisation:

- First of all, minimum and maximum values were added for each angle. These constraints were used to sort out biomechanically infeasible postures. A further advantage was that for a constant sampling rate, the angular resolution was increased because the samples concentrated on the interval of the feasible angles. Therefore, the accuracy could be increased without increasing computational complexity.

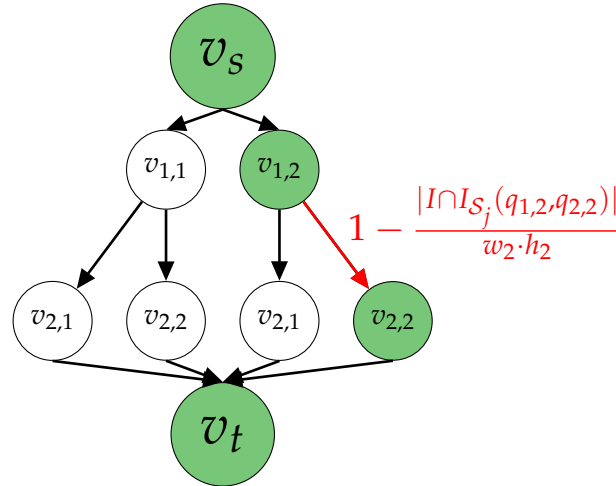
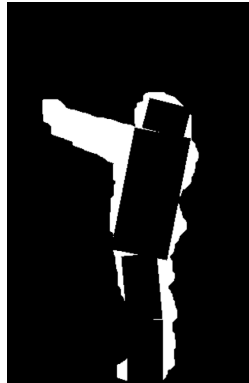


Figure 5.18: Visualisation of the graph built from a solution set with $J = 2$ and $N = 2$. An exemplary path from v_s to v_t is coloured in green.

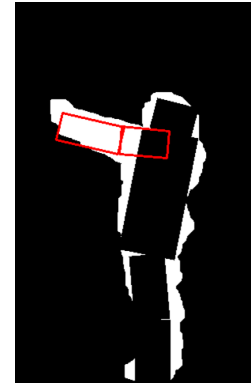
- A hierarchical optimisation scheme was used assuming that the posture of the lower limbs and torso could be distinctively recognised even without knowing the posture of the upper limbs. Thus, the optimisation of the parameters was split into two steps. Having obtained the parameters of \mathcal{M}_{Body} , the algorithm initiated a second optimisation procedure to identify the posture of \mathcal{M}_{Arm} for a fixed \mathcal{M}_{Body} . As preparation of the upper limb optimisation, all body model image pixels, $I_{\mathcal{M}}^{Body}(\mathbf{q}^*) = 1$, in the input image, I , were set to zero. This was done to avoid a solution where the optimised upper limb lies in the body. The preparation workflow is illustrated in Figures 5.19a, 5.19b and 5.19c.



(a) Input silhouette image



(b) Subtracted body model



(c) Fitted upper limb

Figure 5.19: Intermediate results while preparing the fitting of the upper limb model.

This split-up of the optimisation process was advantageous in that the search space was decreased. For example, if there were $N = 10$ samples of each joint for $J = 5$ joints, the search space would be $10^5 = 10000$. In contrast, splitting the search into two steps with $J_1 = 3$ and $J_2 = 2$ would result in a search space of $10^3 + 10^2 = 1100 << 10000$.

- As can be seen in Eq. 5.7, the segment dimensions, w_i and h_i , played an important role in computing the distance. It was important for the localisation quality that these parameters correspond to the antropometric dimensions of the subject. To estimate the parameters, a base model was created first, that had the parameters matching those from the standard, DIN 33402-2 (see Tab. B.1 Appendix). The length and width of each segment were assumed to change, but their ratio among each other should remain similar. When initiating the localisation process, the subject was expected to stand upright. Hence, the height of the subject could be determined from the silhouette image and then the parameters of the base model could be scaled according to that measured height.
- A problem was that the proposed method was susceptible when the upper limbs pointed towards or away from the image plane. As a result of the projection of a 3D object into 2D image space, the 2D length of the arm shortens. The 2D model does not consider this change in terms of segment length. The outcome was that the algorithm computed incorrect angular values. Figure 5.20a shows an exemplary case.

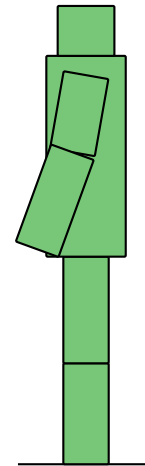
Three approaches were devised to solve this problem. First, a separate method specialised in locating the hand [173] was developed. The method determined an expected torso width and classified all silhouette points lying outside of this width interval as the hand. Afterwards, the distance measure was extended to account for the detected hand in the optimisation task. While the detection of the hand worked well, as soon as they were slightly apart from the body, localising the elbow proved to be challenging. However, without recognising the elbow, it cannot be guaranteed that the algorithm recognised the correct posture. A second approach was to incorporate the segment length as the parameter to optimise. Unfortunately, this led to a set of unforeseen, incorrect solutions - the optimiser had more degrees of freedom in which to find illogical configurations. Besides, there were a variety of cases where adding these parameters dramatically elevated computation time.

The third and final approach produced acceptable results. Upon closer inspection, the observation could be made that every time the shortening of the upper limb took place, the arm in the 2D image was smaller than the model arm. This led to a lower overlap between input image and model arm and resulted in a high distance value, $d_{Approx2D}(I, I_{\mathcal{M}}^{Arm}(\mathbf{q}^*))$. Thus, a standard posture for the arm (see Figure 5.20b) was defined. In case the distance between the model image and image features was above a given threshold, the model parameters were assigned with the standard arm parameters. An advantage of this approach was that it enabled the limb with the highest physical load was tracked. In case both hands were below the shoulder height, the distance measure would be high for the upper limb. Therefore, the algorithm assumed that the upper limb posture was akin to the standard configuration. From an ergonomic assessment point of view, the standard arm and the current configuration would then be equal. In case the upper limb was located above shoulder level reaching forward, the optimisation algorithm could find a small distance mea-

sure. Thus, the optimised configuration was then employed instead of the standard configuration.



(a) Example of the shortened 2D length of the upper limb



(b) Standard arm posture

Figure 5.20: The problem of the shortened upper limb and the most successful solution.

Figure 5.23a summarises the 2D joint localisation workflow.

5.5.4 3D Joint Localisation

As described in Sec. 5.5.3, the disadvantage of the 2D localisation algorithm was that the model assumes a fixed 2D segment length and ignores variations based on the projection of the 3D limbs. The localisation method assumes that the subject is always located at the same distance to the camera being observed from the side perspective. As such, a 3D joint localisation approach was developed to overcome these limitations. The 3D method employed the analysis-by-synthesis approach and integrated the work of Oikonomidis et al. [152] with task specific modifications.

Image Features

The image data came from the depth image created by the Kinect sensor. The masked depth image obtained according to Sec. 5.5.1 was used. The top-left panel of Figure 5.22 illustrates an example of such a masked depth image.

Model

The model \mathcal{M} consisted of a 3D kinematic tree with joint angle ranges corresponding to the degrees of freedom of a human body. Ellipsoid wrappers were modelled around each segment of the skeleton (see Fig. 5.14 for a schematic example). The advantage of

this shape was that there is an analytical representation. The creation of the depth image from the model could be expressed in a closed-form solution. This enabled computing the exact depth value in arbitrarily sized images.

Model Image

The model image, $I_M(\mathbf{q})$, was produced by ray-casting. Ray-casting is a commonly used method in computer graphics to artificially create an image through imitating the underlying physical process.

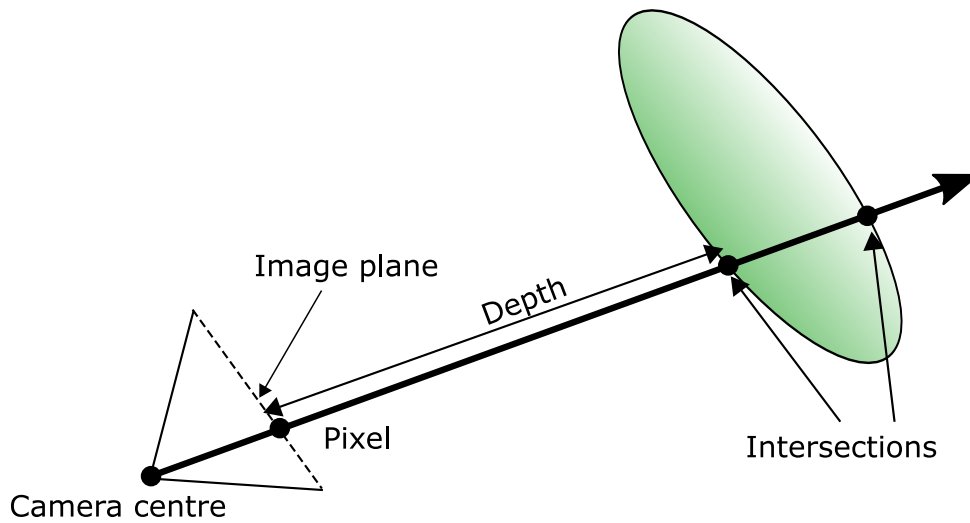


Figure 5.21: Visualisation of the ray-casting approach.

Rays from the camera centre are traced through each pixel of the image plane (see Fig. 5.21). Their intersection point with each wrapper surface is computed. The depth value is the length of the vector from the pixel on the image plane to the nearest point of the intersection. In the following, the formula for computing the depth value for ellipsoid wrappers is derived. First, consider the simplest case - the intersection between a spherical wrapper \mathcal{S} and a ray with the origin e and the direction vector d . The wrapper has a radius of 1 and its centre lies in the origin, $(0,0,0)$. Each point on the ray $\mathbf{x}(t)$ can be described with the vector equation:

$$\mathbf{x}(t) = \mathbf{e} + t\mathbf{d} \quad (5.10)$$

where t denotes a scalar variable. For each point $\mathbf{x}(t)$ on the surface of the sphere, the condition:

$$\mathbf{x}(t) \cdot \mathbf{x}(t) = 1 \quad (5.11)$$

applies, where ‘ \cdot ’ denotes a scalar product. Combining the equations 5.10 and 5.11 results in the intersection point equation:

$$\begin{aligned} \mathbf{x}(t) \cdot \mathbf{x}(t) &= 1 \\ (\mathbf{e} + t\mathbf{d}) \cdot (\mathbf{e} + t\mathbf{d}) &= 1 \\ t^2(\mathbf{d} \cdot \mathbf{d}) + t(2\mathbf{e} \cdot \mathbf{d}) + (\mathbf{e} \cdot \mathbf{e} - 1) &= 0 \end{aligned} \quad (5.12)$$

, where t is the sought-after parameter corresponding to the depth value. Equation 5.12 can be solved using the quadratic formula:

$$t^* = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad (5.13)$$

with $a = \mathbf{d} \cdot \mathbf{d}$, $b = 2\mathbf{e} \cdot \mathbf{d}$ and $c = \mathbf{e} \cdot \mathbf{e} - 1$. If there is no intersection point between the sphere and ray, the term under the square root becomes negative.

Next, the generalised ray-ellipsoid intersection formula is derived. To compute the intersection point of the ray with an arbitrary ellipsoid, the original ray-sphere intersection equation 5.12 had to be modified. Given an ellipsoid wrapper:

$$\mathcal{S} = (\mathbf{c}(\mathbf{q}), \mathbf{R}(\mathbf{q}), \mathbf{S}) \quad (5.14)$$

with the centroid, $\mathbf{c}(\mathbf{q}) = (c_x, c_y, c_z)$, the orientation expressed by the rotation matrix, $\mathbf{R}(\mathbf{q})$, and the extent in the three dimensions expressed by the diagonal scaling matrix, $\mathbf{S} = \text{diag}(s_x, s_y, s_z)$. First, the ray was transformed with:

$$\begin{aligned} \mathbf{e}^* &= \mathbf{S}^{-1}\mathbf{R}^{-1}(\mathbf{e} - \mathbf{c}) \\ \mathbf{d}^* &= \mathbf{S}^{-1}\mathbf{R}^{-1}\mathbf{d} \end{aligned} \quad (5.15)$$

Thereafter, the intersection point was computed by solving Eq. 5.12 with the transformed ray parameters, \mathbf{e}^* and \mathbf{d}^* . To apply the algorithm on a Kinect image, the intrinsic parameters were incorporated into the formula. The device-specific direction vector, d , had to be determined for each pixel. A depth image was initially created with a fixed-depth value for all pixels and converted the pixel coordinates into global coordinates of the Kinect system. The Kinect coordinate system assumed the depth camera centre to be at the coordinate origin. The axis orientations were characterised according to Fig. 5.9. The coordinate conversion was implemented in the driver library [170]. Then, the global coordinates were interpreted as direction vectors and normalised to have a length of 1. The resulting vectors represented d . To obtain a full depth image, the generation process had to be completed for every pixel and wrapper. First, e and d need to be transformed into e^* and d^* for every wrapper. Subsequently, eqs. 5.12 have to be solved. Finally, the smallest t among each wrapper has to be determined. To conclude, the number of intersection computations is the number of pixels times the number of wrappers.

Distance Measure

The distance measure between the images is a weighted sum of two terms:

$$d_{3D}(I, I_{\mathcal{M}}) = c_1 d_{Depth}(I, I_{\mathcal{M}}) + c_2 d_{Overlap}(I, I_{\mathcal{M}}) \quad (5.16)$$

d_{Depth} signifies the depth difference between model image and masked depth image and is defined as

$$d_{Depth}(I, I_{\mathcal{M}}) = \sum_{x,y} f_1(x,y) \min(|I(x,y) - I_{\mathcal{M}}(x,y)|, D_{Max})$$

$$\text{with } f_1(x,y) = \begin{cases} 1, & \text{if } I(x,y) > 0 \wedge I_{\mathcal{M}}(x,y) > 0, \\ 0, & \text{else.} \end{cases} \quad (5.17)$$

Oikonomidis et al. [152] contended it was important to clip the depth distance in each pixel as soon as it exceeds a threshold, D_{Max} , to avoid outliers. $d_{Overlap}$ is the number of pixels where only one image had a depth value:

$$d_{Overlap} = \sum_{x,y} f_2(x,y)$$

$$\text{with } f_2(x,y) = \begin{cases} 1, & \text{if } I(x,y) = 0 \wedge I_{\mathcal{M}}(x,y) > 0, \\ 1, & \text{if } I(x,y) > 0 \wedge I_{\mathcal{M}}(x,y) = 0, \\ 0, & \text{else.} \end{cases} \quad (5.18)$$

It is similar to the overlap distance measure of the 2D joint localisation algorithm.

Optimisation

Similar to the 2D algorithm, there was the option to choose between different optimisation strategies. At the outset, discrete optimisation was eliminated. Contrary to the 2D problem, roughly 60 parameters had to be optimised. Even with elaborated algorithms, this computation time became unacceptably high as artificial depth image generation is more computationally expensive than 2D silhouette generation. Furthermore, derivative-based approaches, such as gradient descent methods, could not be used. In that case, the derivative of the objective function (Eq. 5.16) with reference to the parameters had to be computed. As the equations, 5.17 and 5.18, are piecewise functions that include binary conditions, it could not be guaranteed that the objective function would be differentiable in all domains, and this is the reason this option was discarded. Following [123], the PSO was applied. This stochastic optimisation method was introduced by Eberhart and Shi [174]. Contrary to derivative-based optimisation methods, the algorithm does not compute the derivative of the distance measure. PSO evaluates multiple hypotheses and computes the search direction through the interaction of the hypotheses.

The basic principle behind this is that p hypotheses, $\mathbf{q}_{t,1}, \dots, \mathbf{q}_{t,p}$, in the search space, which are called particles, are iteratively moved towards the final solution. In each itera-

tion, t , the particles move in the direction of the velocity vector, $\mathbf{v}_{t,i}$:

$$\mathbf{q}_{t+1,i} = \mathbf{q}_{t,i} + \mathbf{v}_{t,i}. \quad (5.19)$$

The velocity vector is a weighted sum of three components:

$$v_{t,i} = c_{social} \cdot r_1 \cdot (\mathbf{q}_{global} - \mathbf{q}_{t,i}) + c_{cognitive} \cdot r_2 \cdot (\mathbf{q}_{local,i} - \mathbf{q}_{t,i}) + c_{inertial} \cdot \mathbf{v}_{t-1,i} \quad (5.20)$$

The first component is the movement towards the currently best particle, \mathbf{q}_{global} . It is called the social component. The second is the movement of a particle towards its own best position ever achieved $\mathbf{q}_{local,i}$. The literature sometimes refers to it as the cognitive component. The last component is the velocity vector of the last iteration, the inertial component. The variables, r_1 and r_2 , are random numbers between 0 and 1. In their experiments on standard optimisation benchmarks, Bastos et al. [175] put forth that without using random numbers, the PSO operates significantly worse.

To improve the performance for the particular task, several modifications to the core algorithm were introduced:

- All components of the search space were normalised to 0 – 1. The particles contained translational and rotational parameters with different units (mm and rad) and magnitudes. The normalisation resulted in appropriate scaling of each parameter's movement.
- While updating the particles, they may leave the solution space. The "damping" method [176] was used to reset them into the solution space.
- The temporal coherence of the video stream was leveraged by uniformly initialising the particles in the first iteration, $\mathbf{q}_{1,i}$, with values around the determined parameters, \mathbf{q}_{global} , from the last depth frame.
- Similar to the 2D approach (see Sec. 5.5.3), a two-step hierarchical search space decomposition scheme was applied. During the first step, the model was optimised by holding the upper limb parameters fixed. With the second step, the torso and lower limb parameters were fixed and the upper limb parameters were optimised. The number of variables to be optimised and, thus the search space, became smaller and decreased the likelihood of running into local minima.

Figure 5.22 depicts the result of the 3D joint localisation algorithm for one frame. The qualitative result shows that the ellipsoid wrapper shape is not an optimal approximation of the human body. Poppe [128] lists alternative wrapper shapes used in literature, such as cylinders, quadrics or meshes of human models. In the scope of this thesis, these alternatives have not been tested due to two reasons: Firstly, changing the wrapper shape requires altering the model image generation process. In order to achieve acceptable computation time, the rendering process has to be parallelised on the GPU. The parallel implementation for each alternative is challenging and tedious. As an example, using meshes involved a reimplementation of the whole PSO routine [177]. Secondly, the potential benefit does not justify the required implementation effort. Using a different wrapper type

does not necessarily improve the quality or robustness of the joint localisation. A better fitting shape might achieve a lower distance measure value, however, does not guarantee more accurate model parameters. Thus, the difficulty of the optimisation problem will likely remain despite the use of better approximations of the human body.

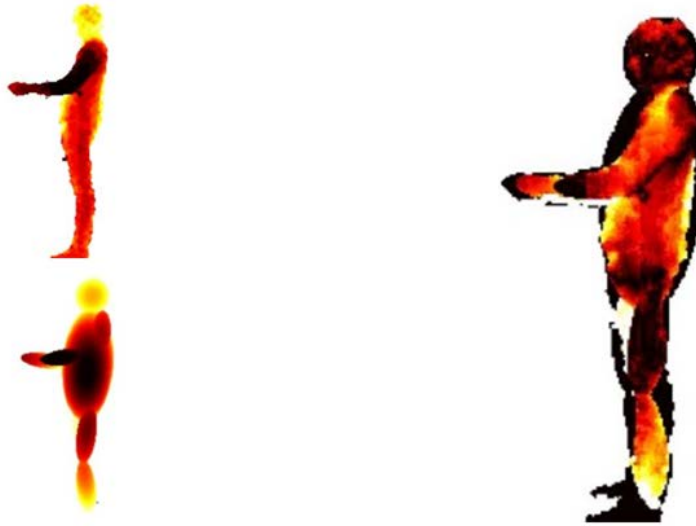


Figure 5.22: Exemplary result of the 3D joint localisation algorithm.

Summing up, the overall workflow is portrayed in Fig. 5.23b.

5.6 Posture Classification

Having obtained the model parameters for the current frame, they were used to classify the posture. For the 2D case, the joint angles were directly employed as inputs for the classifiers. In the 3D situation, joint orientations were added to the feature vector.

A variety of classifiers were used to test the classification algorithm. The difference between the classifiers was rooted in the way they represented the learned knowledge:

- The K-Nearest Neighbour (KNN) classifiers represent an instance-based approach where a new data point is compared with all data points of the training dataset according to a pre-defined similarity metric. The euclidean distance was used as the dimensions are bestowed with an equal range. The data point is assigned to the label of the most similar data point in the training dataset.
- The Support Vector Machine (SVM) with linear kernel (LIN-SVM) is a method that symbolises the knowledge as linear decision boundaries between the classes. Based on the small number of trainable parameters compared to more sophisticated methods, linear classifiers require only a small dataset and are less prone to overfitting.

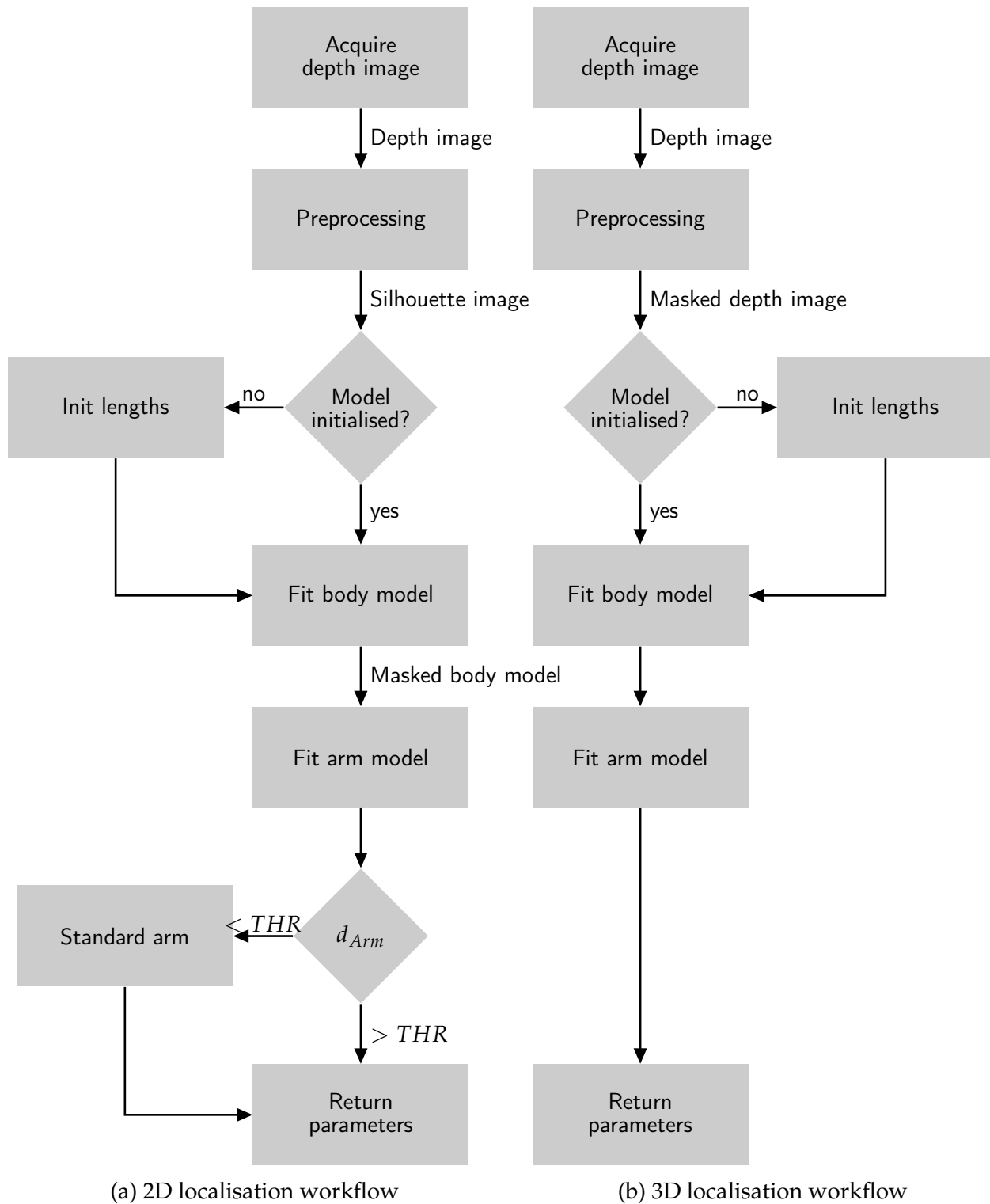


Figure 5.23: Flow diagram of 2D and 3D joint localisation approach.

- The SVM with radial basis kernel (RBF-SVM) also represents the knowledge by decision boundaries. However, the boundary is not necessarily linear, enhancing the expressiveness of the model, but also the likelihood to overfit the training data.

- The Random Forest (RF) does not explicitly rely on distance measures. Rather than comparing the whole data points, it selects relevant features that best separate the dataset and identify interesting value ranges. Shotton et al. [11] showed they are among the best performing classifiers in recent years.

5.7 Conclusions

This chapter presented the pipeline of the automated vision-based posture classification component, the core of the posture assessment component. The central task was to recognise the worker's posture in the image. Based on the minimal spatial distance requirement of the EAWS, the conclusion was made that a camera observing the worker from the side at a distance of around 3.5m suited the task.

A literature research uncovered that using joint angle features was appropriate for the task as the joint position coordinates to be extracted were needed for the posture optimisation, as well. Determining the joint angle values necessitated a joint localisation algorithm. The generative localisation approach was favoured as it needs no training data and has been the method of choice over the last years.

Two joint localisation approaches were proposed to attain the joint positions. Both are unified by the generative framework, which makes them easily modifiable. A quick 2D approach and a slower approach able to obtain 3D coordinates of the joints have been put forth. In their base versions, both approaches have limitations concerning robustness and runtime. Hence, slight modifications were introduced to resolve these issues. The major challenge was to deal with changing 2D lengths of the segments because of projections onto the image plane.

Having obtained the joint positions, computing the joint angles and applying classifying algorithms is straight-forward. Thereby, multiple classification algorithms (KNN, LIN-SVM, RBF-SVM, RF) were proposed to be tested.

6 Posture optimisation

This chapter discusses the posture optimisation component of the WPC, which comes into play when the result of the posture assessment algorithm indicates the need to re-adjust the workplace. The aim of the posture optimisation method is to find a pose for the workpiece that permits accomplishing the task at hand in an ergonomically superior posture.

6.1 Problem Statement

This chapter addresses the research question, RQ3:

How can an automated and adaptive in-process workplace design be realised on various types of actuators?

The algorithm takes the human model obtained from Sec. 5.5 including estimated joint angles and segment lengths as input (see Fig. 6.1a). The fitted model was used to identify the characteristics of the actual task the user intends to accomplish. The posture optimisation algorithm finds an alternative workpiece pose, which enables the worker to maintain a more health-preserving posture while accomplishing this reference task. In brief, the algorithm outputs the necessary actuator configuration and shows what this "ergonomically superior posture" resembles (see Figure 6.1b). Figure 6.1c exhibits an exemplary situation with input and output postures visualised.

The difficulty of this task is that the posture the human worker will adopt after a hypothetical workplace adjustment is not known beforehand - the system has to predict this working posture. Furthermore, an approach is necessary which searches for the workpiece pose, resulting in an ergonomically preferred posture. An alternative approach to searching for human posture and workpiece pose simultaneously is to fix the human posture to an ergonomically preferred one and only search for the corresponding actuator parameters. This procedure can work for some cases. However, there is no backup solution when an appropriate adjustment is not feasible for the given posture. When the algorithm has the possibility to slightly adjust the worker's posture, it can find backup solutions in case the optimal posture cannot be realised. This is why adjusting worker posture as well as parameters is preferred.

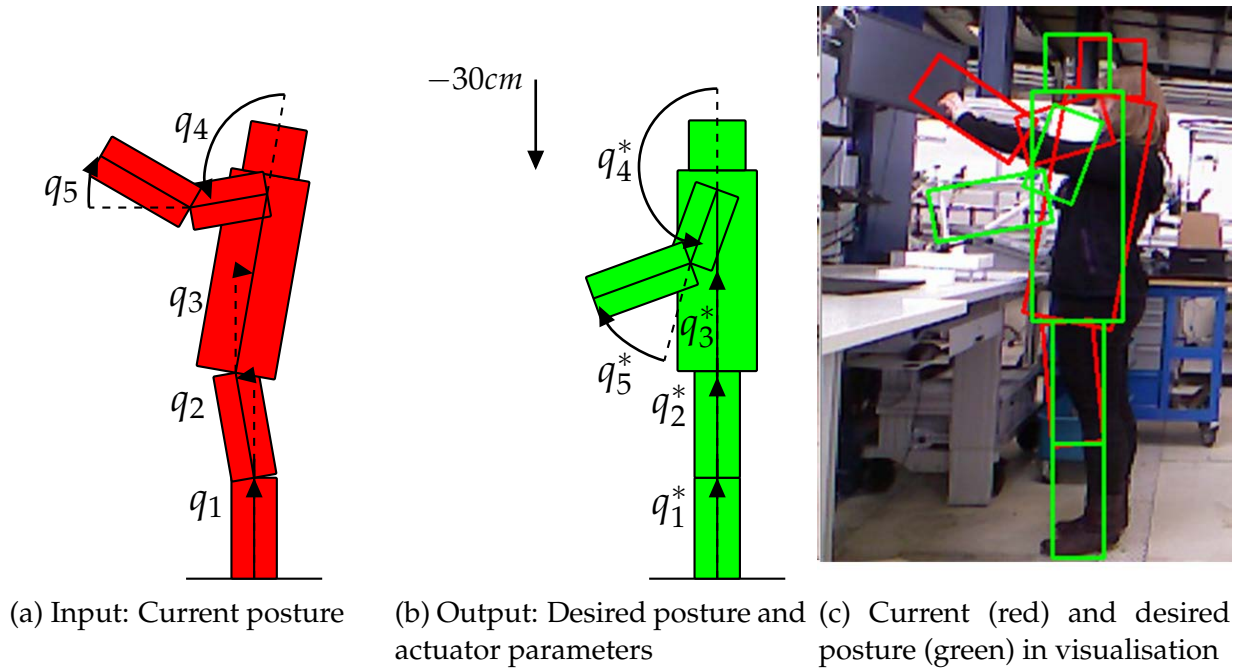


Figure 6.1: Input and output of the posture optimisation algorithm for the example of a height-adjustable actuator.

In order to properly avoid physical load, the developed method must be able to provide a feedback within a few seconds. One additional requirement was that the algorithm considers the capabilities and available DoF of the actuator employed. Apart from optimising the workplace for one posture, the dynamic behaviour of the WPC has to be taken into consideration. While frequent adjustments of the workpiece pose can minimise the postural load, they also introduce annoying interruptions that impede the process' efficiency. The optimisation criteria, "ergonomics" and "flexibility", contradict each other. Hence, the task was to develop a control strategy that outputs when a computed workplace adjustment is needed. The control strategy strives to determine an acceptable compromise between postural load and adjustment time. Figure 6.2 summarises the research questions to be answered in this chapter.

6.2 Design Alternatives

6.2.1 Posture Prediction

Posture prediction methods output a hypothesised posture given an input 3D point, which the human being touches [80]. This task is especially prominent in the field of DHMs, where it is used to determine ergonomic parameters through simulation. The approaches proposed in the previous literature can be clustered into two groups - optimisation-based methods [178, 179, 180, 80, 181, 182, 60] and learning-based methods

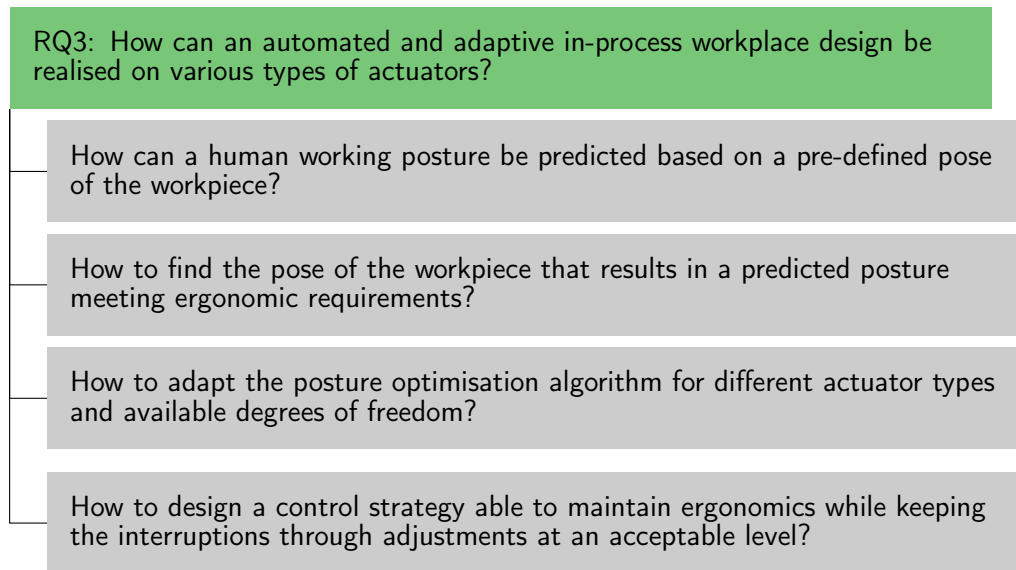


Figure 6.2: Summary of sub-research questions RQ3 implies.

[183, 184, 185, 186]. Optimisation-based approaches establish the predicted posture as the solution of an optimisation problem encoding the criteria and conditions for most likely postures. The optimisation problem has to be manually modelled. Human engineers and researchers are required to design the objective function and the constraints. In basic terms, the behaviour has to be translated into mathematical terms. Fortunately, the ergonomic guidelines provide heuristics. As an example, the EAWS directly defines the most preferable posture as it assigns the lowest score to it. Another advantage is that data generation is not necessary. Finally, as the mathematical formulation of the objective function is already present, the search for an appropriate workpiece pose can be realised with a variety of standard optimisers. The extension from posture prediction into a workpiece pose search is natural.

As the modelling of realistic human behaviour with objective functions is challenging, learning-based methods have been proposed. Instead of modelling the constraints and objective functions by hand, the system is provided a set of exemplary workpiece poses and the corresponding posture. The posture prediction algorithm then learns the mapping between workpiece pose and posture. Learning-based methods appear to be easier to use. The system behaviour is modelled by examples instead of encoding the knowledge into mathematical formulae. Moreover, with optimisation-based approaches, the designer has to work with all the objective and constraint parameters, such as the weighting of different criteria. If the number of parameters are high, manual modelling becomes tedious. In learning-based methods, the system automatically deals with these parameters. Therefore, they can be addressed with more complex behaviour with satisfactory effort. Although diminishing the effort for modelling, the advantage of learning-based methods is offset by the required amount of data. The data generation process is tedious. The space of possible workpiece poses has to be densely sampled and the corresponding joint angle values of the worker have to be determined. In addition, a variety of subjects

with different anthropometrics have to be considered. Having learned the postural mapping, the options for searching for the best workpiece pose are limited. As a result of the learned mapping not necessarily being explicitly expressed in a mathematical function, only black-box optimisation approaches remain the viable option.

Selected Approach

The optimisation-based approach was selected. Initially, there are existing ergonomic rules used to design the objective function. Hence, the modelling effort for the given case was reduced. Moreover, the tedious data generation process becomes obsolete. Still, reference posture examples can be seamlessly integrated into the approach. Table 6.1 outlines the criteria employed to choose a posture prediction approach.

Table 6.1: Overview of posture prediction alternatives. Columns highlighted in green represent the selected approach.

Criterion	Learning-Based	Optimisation-Based
Training data required	yes	no
Computational complexity	low	low
Incorporate ergonomic knowledge	by training examples	by mathematical modelling or training examples

Risk Analysis

The optimisation-based approach requires each decision criterion to be stated by the means of formal language. However, there might be essential ergonomic guidelines, which are too complex to be expressed by a formula leading to an incomplete decision model. Furthermore, as the posture prediction problem is presumably non-linear, there is no guarantee that an optimiser will always terminate in a desired local minimum.

6.2.2 Workpiece Pose Search and Workplace Design

The task of the WPC is to discover a workpiece pose that would permit the worker to perform the task at hand in a comfortable posture. This problem has not been intensively treated in the literature. The only publication found was Mainprice et al. [13] who developed a method to plan the best path to hand over an object to a human being. Therein, ergonomic criteria were included in the robot's handing over posture. However, the objective used to guide the posture of the end effector was rather simple - it encompassed just the distance to the hands and visibility criteria. Postural ergonomics was not considered (see Fig. 6.3).

Strongly related to the present task is the procedure of designing a workplace layout. Hence, this field was investigated attempting to find a solution that could be modified for the original workpiece pose search task. Workplace layout design adheres to a more

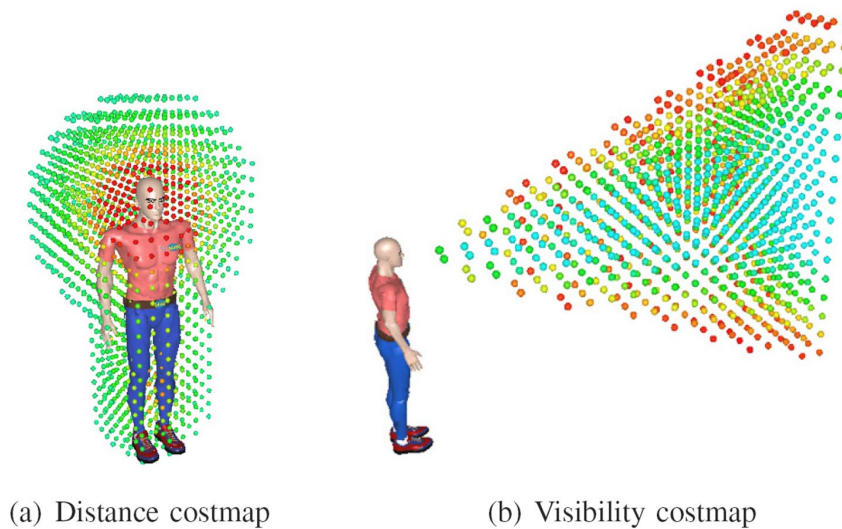


Figure 6.3: Mainprice et al. [13] used an optimisation based approach, however, the objective function does not consider the posture of the subject.

or less standardised procedure, which is sometimes referred to as methods engineering [187] in literature. The workflow consists of five steps [19]: project definition, data gathering and analysis, formulation of alternative solutions and methods, evaluation and refinement of the best alternative according to the evaluation (see Fig. 6.4).

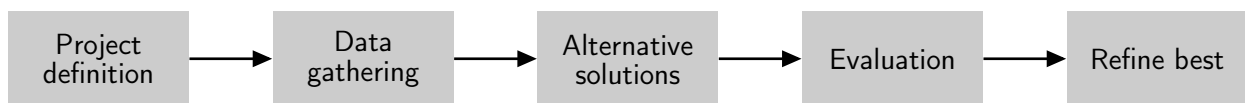


Figure 6.4: Methods engineering workflow.

As would be obvious, the project or the target has to be defined. In the present case, this could be designing the workplace geometry such that awkward back postures do not need to be adopted to accomplish any tasks. Afterwards, data from the existing or initial solution needs to be obtained. In the present case, this would be an ergonomic assessment of the processes executed at the target workplace. Next, one of the most work-intensive steps has to be carried out - alternative methods or workplace layouts have to be designed. Control variables (factors), like height of the workplace or height of the chair or placement of a particular box, are selected and their possible values are formulated. Given n control variables with m possible values each, there are theoretically m^n solutions to be appraised. This factorial design becomes exponentially work intensive in relation to the factor size - if factor values are continuous, discretisation or regression approaches [19] have to be applied. The problem worsens as the criteria for evaluation are highly complex. In terms of workplace layout, there are many criteria to be considered. The criteria to be optimised are often contradictory, such as safety and efficiency [188]. As well, layouts must be tested with a variety of anthropometrics [7]. Building physical prototypes is expensive, which is why DHMs have been widely used. Finally, the most promising alternative has to be designed in detail. The process is not necessarily linear

but the steps might be repeated multiple times. For example, evaluations and refinements of the alternatives can be conducted in an iterative manner [189]. Concerning the task of finding a workplace layout design to meet ergonomic goals, the literature has presented four approaches: selection out of manually chosen alternatives [61, 64, 53], expert systems [190], brute-force search [60] or non-linear optimisation [8].

Selected Approach

Manually chosen alternatives require a human operator to select the possible solutions. As mentioned earlier, a factorial design requires the evaluation of a large number of alternatives. To speed up the design process, specialised heuristics have been applied [189]. Although reducing the search space, the proposed method still contains a remarkable proportion of brute-force search. The efficiency problem also arises with brute-force approaches, where the computer evaluates all solutions. As the number of options grows exponentially with the number of factors and parameters, brute-force search is only viable if the search space is small. Expert systems apply a set of rules to find a solution for the given problem, thus the search for a solution is efficient. However, the system needs a tremendous human effort to create data and the set of rules. Finally, there are optimisation-based approaches. As described in Sec. 6.2.1, the posture prediction problem can be naturally extended to the workpiece pose search problem and effectively resolved with standard non-linear optimisation algorithms. As the optimisation-based approach was used for posture prediction, it would be straightforward to further follow this path for the workpiece pose search. Table 6.2 shows the criteria applied to determine a workpiece pose search approach.

Table 6.2: Overview of workpiece pose search alternatives. Columns highlighted in green represent the selected approach.

Criterion	Manual	Brute-force	Expert system	Optimisation
Human intervention required	yes	no	no	no
Manual development effort	high	low	high	low
Computational complexity	low	high	low	low

Risk Analysis

Similar to the risks of the selected posture prediction approach, a central problem will be whether the proposed approach is sufficiently robust against becoming trapped into undesired local minima during the optimisation process.

6.2.3 Control Strategy

Consider a process where the worker has to perform tasks on different regions of the workpiece. If the workpiece dimensions are longer than the length of an arm, the worker

probably has to adopt different postures from kneeling down to reaching up in order to accomplish each step. Each time the posture has changed, a re-adjustment of the actuator is proposed. Especially when the worker is moving, there are frequent adjustments - the posture changes in each frame. Even when the worker remains in a completely static posture, the tracking algorithm might introduce slight inaccuracies leading to a jittering of the estimated body parameters, which again result in adjustment proposals.

A system that proposes a new actuator configuration for every frame introduces an unacceptable amount of workflow interruptions. Most of these adjustments might be unnecessary as the time that remains in the awkward posture is low. Their ergonomic improvement could also be negligible. Further, slight adjustments when the worker's current posture is near the optimum would introduce more costs than benefits. Frequent false alarms lead to the cry wolf effect [191] - workers mostly ignore the proposals. Hence, the expected effect on health is lost.

The WPC has to recognise the most relevant ergonomically critical situations and only propose an adjustment then. The problem is that with recent ergonomic assessment tools, the criticality of a process can only be found out after the process has ended. For example, the EAWS refers to the accumulated physical load after all tasks have been accomplished. The decision whether to initiate an adjustment, however, has to be made during the process. The research field most related to the present task is optimal control. Optimal control methods attempt to determine a control strategy for a dynamic system that optimises the performance index represented by L and ϕ :

$$\begin{aligned} \arg \min_{u(t)} \int_0^{t_{End}} L(x(t), u(t)) dt + \phi(x(t_{End})) \\ \text{subject to} \end{aligned} \quad (6.1)$$

$$\begin{aligned} \dot{x}(t) &= f(u(t), x(t)) \\ x(0) &= x_0 \end{aligned}$$

The functions, $x(t)$ and $u(t)$, indicate a state of the system and a control function, respectively. The performance index is represented by L , which denotes the cost for the particular control function, $u(t)$, and ϕ , which signifies the cost of the resulting system state at a particular time, t_{End} . The general optimal control task is to minimise this functional subject to the state equation, $\dot{x}(t)$, which is indicative of the system behaviour and a given initial state, x_0 .

Selected Approach

Translating the general definition of the optimal control problem to the present task results in the following instantiations:

- $x(t)$ contains the joint angles of the worker, \mathbf{q} , and the EAWS score at time t
- $u(t)$ denotes the actuator configuration, \mathbf{p} , at time t
- L signifies the path of the actuator, $\int u(t) dt$

- the terminal cost, ϕ , refers to the EAWS score after the process
- $\dot{x}(t)$ represents the scheme to compute the joint angles after the workplace layout has been adjusted through actuator configuration, $u(t)$

Optimal control methods normally require at least a computable gradient of the optimisation problem [192]. However, as a result of the complexity of the state equation, $\dot{x}(t)$, this is not possible. Computing $\dot{x}(t)$ necessitates predicting the joint angles after adjustment, $u(t)$, which is a highly non-linear function. It is not even analytically expressible (see Sec. 6.2.1). In order to continue to find a reasonable control strategy, the structure of the EAWS was exploited deriving simple control rules to reduce the ergonomic load as well as the actuator activity.

Risk Analysis

Finding an optimal control strategy with reference to a particular task normally requires knowing all occurring postures. However, when operating in-process, complete information is not available. Hence, the control strategy has to base its decisions on data from the present and the past. This problem of incomplete information might impair the ability of the control strategy to find a fair compromise between ergonomic soundness and flexibility.

6.3 Optimisation-Based Approach

The optimisation-based approach was chosen to solve the posture prediction and workpiece pose optimisation task. This section discusses how to mathematically state the optimisation problem. The mathematical objectives and constraints were partly derived from the EAWS guidelines as well as from own assumptions of the task. Although there is no "the" ideal posture to be adopted, there are a set of guidelines in the ergonomics literature describing how "preferable" postures and workpiece poses look. For example, Schmidtke and Jastrzebska-Fraczek [95] approximated preferred working heights for particular tasks. Moreover, the EAWS [34] scoring scheme leads to the observation that the postures "standing and walking in alternation" and "standing with body support" (see Fig. 6.6) yield the lowest scores. Thus, they are considered the most preferred postures. Similarly, other observational ergonomic assessment tools, including REBA [29], define what they interpret as the "ideal posture". The ideal posture also depends on the postures adopted beforehand: Neuhaus et al. [51] stated that a worker shall only sit for 50% of the time and stand for the remainder. These examples suggest that although being able to properly translate the guidelines into formulae, this does not guarantee that the system always finds the "best" solution - the "ideal posture" is ambiguously defined.

The idea of an ideal posture is simplified by using the EAWS posture with the lowest score. If there is another posture better suited as an ideal posture, the optimisation prob-

lem could be adapted without effort. When necessary, the mathematical modelling allows multiple ideal postures. The rest of this chapter is structured as follows. First, the general optimisation-based framework for posture prediction is described. Then, objectives and constraints are specified and their mathematical properties are analysed. Next, the posture prediction problem is extended into a workpiece pose search problem and the algorithm to solve it is outlined. Finally, the dynamic behaviour of such a workpiece pose selection algorithm is evaluated and an approach to diminish the amount and time of interruptions through adjustment is presented.

The optimisation-based approach models the choice of the worker's posture by an objective function, $f(\mathbf{q}, \mathbf{p})$. This function takes the posture parameters, \mathbf{q} , and the actuator parameters, \mathbf{p} , as inputs and returns high values for unlikely postures and low values for likely postures. Additionally, there can be equality constraints, such as $g_i(\mathbf{q}, \mathbf{p}) = 0$, or inequality constraints, such as $h_j(\mathbf{q}, \mathbf{p}) \leq 0$, which limit the space of feasible postures and actuator configurations. The solution to the possibly non-linear and non-convex problem is:

$$\begin{aligned} \mathbf{q}^*, \mathbf{p}^* &= \arg \min_{\mathbf{q}, \mathbf{p}} f(\mathbf{q}, \mathbf{p}) \\ &\text{subject to} \\ &g_i(\mathbf{q}, \mathbf{p}) = 0 \\ &h_j(\mathbf{q}, \mathbf{p}) \leq 0 \end{aligned} \tag{6.2}$$

where \mathbf{p}^* represents the actuator adjustment and \mathbf{q}^* the predicted posture.

In the following, the definition for a segment $\mathcal{S}_j = (\mathbf{q}_j, w_j, h_j)$ from Eq. 5.4 is used. The posture optimisation algorithm takes the tracked posture, \mathbf{q} , and the segment lengths, h_j , as input. A visualisation of \mathbf{q} is shown in Fig. 5.16b. The values, \mathbf{q}_i , represent the angle between a segment (straight line) and its predecessor in the kinematic chain (dashed line). The angle, \mathbf{q}_1 , represents the angle between the foot segment and the global y axis. Note that \mathbf{q}_i can be positive as well as negative depending on which side of the local y axis the current segment lies in. There are two body landmarks that play an essential role: 1) there is the 2D location of the foot, which in the following is notated as \mathbf{x}_{Foot} ; and 2) there is the position of the hand, \mathbf{x}_{Hand} (see Fig. 6.5a). In order to model the human as a kinematic chain instead of a tree, the assumption was made that hands and feet are aligned such that left and right limbs represent the same 2D point. Given the 2D point of the foot, \mathbf{x}_{Foot} , the hand position, \mathbf{x}_{Hand} , could be calculated using 2D forward kinematics:

$$\begin{aligned} \mathbf{x}_{Hand}(\mathbf{q}) &= \mathbf{x}_{Foot} + \sum_{j=1}^J \mathbf{R} \left(\sum_{i=0}^j \mathbf{q}_i \right) * h_j \\ &\text{with} \\ \mathbf{R}(\mathbf{q}_i) &= \begin{pmatrix} \cos \mathbf{q}_i & -\sin \mathbf{q}_i \\ \sin \mathbf{q}_i & \cos \mathbf{q}_i \end{pmatrix}. \end{aligned} \tag{6.3}$$

Moreover, the term “working direction”, \mathbf{d}_w , is introduced. In Enomoto et al. [60], this vector is referred to as the assembly motion vector. A visualisation of \mathbf{d}_w can be found in Fig. 6.5b. The working direction, \mathbf{d}_w , denotes the direction the forearm of the worker is pointing at. It represents the direction of the assembly motion. Note that \mathbf{d}_w does not provide information on the exact pose of the wrist - it is only an approximation. Often, \mathbf{d}_w will not be used in vector representation, but as $\Phi(\mathbf{d}_w)$, the angle in reference to the global y axis. $\Phi(\mathbf{d}_w)$ is computed as:

$$\Phi(\mathbf{d}_w) = \sum_j \mathbf{q}_j \quad (6.4)$$

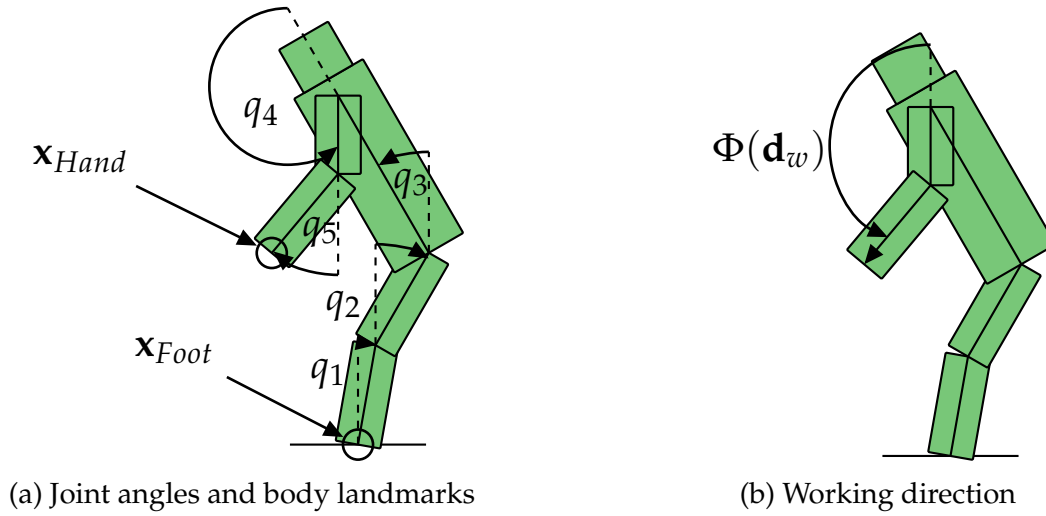


Figure 6.5: Notations used for describing the details of the posture optimisation approach.

6.4 Optimisation Objectives

This section describes the specific appearance of the objective function, $f(\mathbf{q}, \mathbf{p})$. In fact, human behaviour cannot be modelled by one goal, but requires stating multiple, possibly competing goals [180]. Apart from the ergonomic soundness of the posture, the optimisation has to make certain that the task can be properly executed. Thus, a Multiple Objective Optimisation (MOO) problem had to be stated. Modelling and solving posture prediction as MOO problems was previously investigated by Yang et al. [179] and Marler et al. [180]. Both observed that stating the problem as MOO leads to more realistic postures than independently optimising each objective. Their studies proposed three schemes to approach a MOO. These are namely “weighted sum”, “min-max method” and “global criterion”. However, in their experimental results, they stated that the postures computed only marginally differed. In this work, the weighted sum method was elected as it was considered as the easiest to implement. In order to transfer this multi-objective optimisation problem into a problem solvable by single-objective solvers, $f(\mathbf{q}, \mathbf{p})$ was modelled as a weighted sum of particular objective functions, $f_i(\mathbf{q}, \mathbf{p})$. The final objective

function could then be acquired by computing the weighted sum of the partial goals:

$$f(\mathbf{q}, \mathbf{p}) = \sum_i c_i f_i(\mathbf{q}, \mathbf{p}) \text{ with } \sum_i c_i = 1. \quad (6.5)$$

Changing the weights, c_i , yield different pareto-optimal solutions. This thesis concentrates on determining one pareto-optimal solution, leaving the weighting task to the workplace designer.

6.4.1 Posture Objective

The posture objective models the goal of the human being to adapt a posture considered comfortable and “ergonomic”. In human factors literature, this posture is occasionally referred to as a “neutral” posture. The posture objective is independent from the actuator configuration as it only analyses the predicted posture. As stated in Sec. 6.3, there is no definition of an ideal posture. Since the EAWS was employed to assess the postures, the ideal posture was defined as that with the lowest EAWS points (see Fig. 6.6 for visualisation). Hence, the objective function was the distance of the current posture to the ideal posture, \mathbf{q}_{ideal} :

$$f_{Posture}(\mathbf{q}) = \|\mathbf{q} - \mathbf{q}_{ideal}\|^2. \quad (6.6)$$

There are other options for designing a distance function between the current posture and the ideal one, such as the absolute value, $\|\mathbf{q} - \mathbf{q}_{ideal}\|$. Quadratic terms have the advantage that they are twice differentiable over the whole space of \mathbf{q} , which is not the case for absolute functions. This mathematical property was useful when applying optimisation algorithms that utilise function derivatives to establish the search direction. Another property of the function is that a large deviation of one joint angle is penalised more severely than several small deviations of all joint angles.

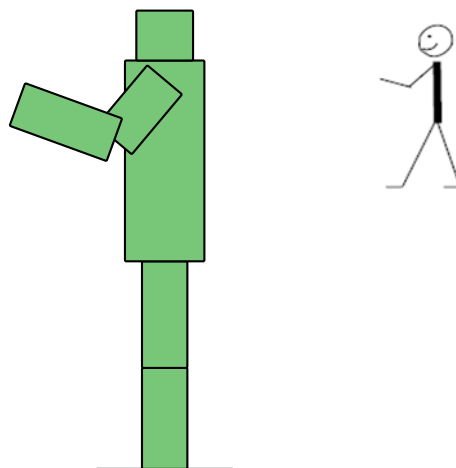


Figure 6.6: Visualisation of the “ideal” posture, \mathbf{q}_{ideal} , defined in the EAWS.

If multiple ideal postures exist, one way to incorporate them into the model is through solving the optimisation problem for different values for \mathbf{q}_{ideal} and choose the solution producing the lowest objective function value.

6.4.2 Centroid Objective

The algorithm presupposes that the subject chooses a posture that would enable them to keep their balance. Similar to the posture objective, the centroid objective only analyses the predicted posture. The stability of a stance was described by the location of the centroid of the model. For simplicity, all segments in the model are assumed to have the same density. Thus, the centroid of a segment only depended on its length. The centroid of the whole body can be computed as:

$$\mathbf{x}_{Centroid}(\mathbf{q}) = \frac{1}{2} \sum_{j=2}^J (\mathbf{x}_j(\mathbf{q}) - \mathbf{x}_{j-1}(\mathbf{q})). \quad (6.7)$$

The higher the horizontal distance of the centroid to the foot, the more weight will be shifted away from the foot, meaning it would be harder to maintain a stable stance (see Fig. 6.7 for visualisation). Therefore, the centroid objective attempts to keep the centroid of the model near the foot in the horizontal direction:

$$f_{Centroid}(\mathbf{q}) = (\mathbf{x}_{Foot,x} - \mathbf{x}_{Centroid,x}(\mathbf{q}))^2. \quad (6.8)$$

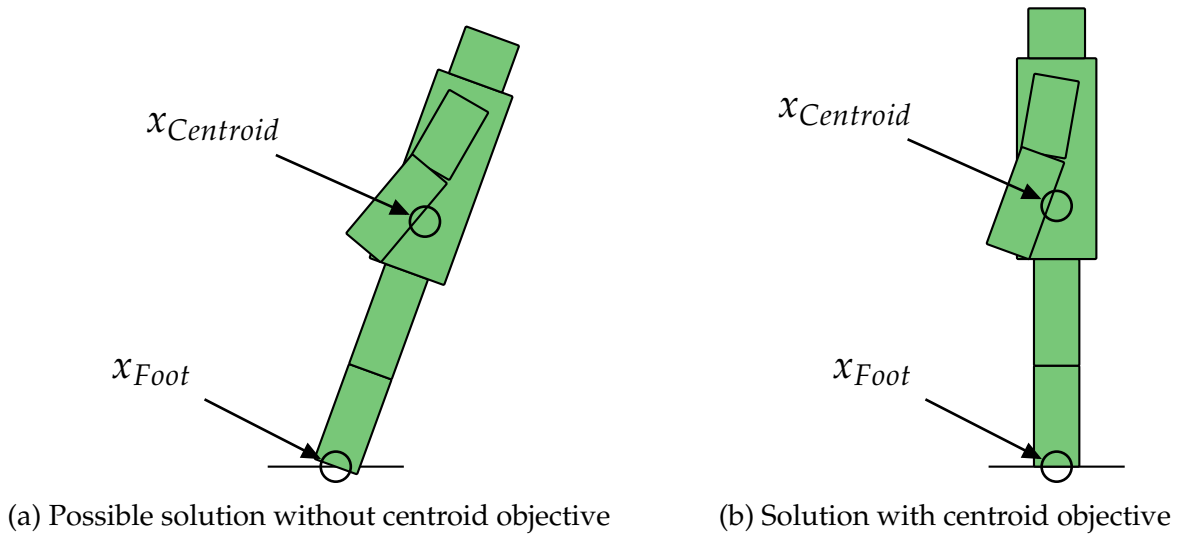


Figure 6.7: Visualisation of the centroid objective.

6.4.3 Working Direction Objective

A proposed adjustment not only must guarantee that the adopted posture is ergonomic. It also needs to ensure that the worker can properly perform the task. The most health-

preserving posture will not be adopted if accomplishing the current task with it is not possible. For example, if the worker intends to drill a hole into the bottom of the workpiece, there is no use in doing so when the top faces the user after adjustment. Hence, the first step before computing the adjustment is to discern which task needs to be accomplished. The task interpretation was derived from the original posture. The spatial relationship between workpiece and working direction were of particular interest. Letting $\phi_{Dir}(\mathbf{p})$ be the orientation of the workpiece encoded as an angle to the global y-axis, the working direction objective is defined as:

$$\begin{aligned} f_{Dir}(\mathbf{q}, \mathbf{p}) &= ((\phi(\mathbf{d}_w) - \phi_{Dir}(\mathbf{p})) - (\phi(\mathbf{d}_{wBefore}) - \phi_{Dir}(\mathbf{p}_{Before})))^2 \\ &= ((\sum_j \mathbf{q}_j - \phi_{Dir}(\mathbf{p})) - (\sum_j \mathbf{q}_{Before,j} - \phi_{Dir}(\mathbf{p}_{Before})))^2 \end{aligned} \quad (6.9)$$

where \mathbf{p}_{Before} and \mathbf{q}_{Before} denote the input actuator and posture parameters. In brief, the working direction objective strives to maintain the same spatial relationship between the workpiece and working direction before as well as after the adjustment of the workplace. If this relationship, for whatever reason, is less critical, it can be weighted less or even left out in extreme cases. As the particular definition of the working direction objective is dependent on the type of the actuator, the objective will be specified in the relevant sections.

6.4.4 Contact Point Objective

Similar to the working direction objective, the contact point objective ensures that the user is able to fulfil the given task after workplace adjustment. In the present case, the reachability of the contact point is the focus. The contact point, $\mathbf{x}_{Contact}$, determines the location on the workpiece where the current operation takes place. Similar to the working direction objective, it is extracted from the input posture. However, as the contact point is given relative to the workpiece, the workpiece dimension is required. The contact point objective is modelled as:

$$f(\mathbf{q}, \mathbf{p}) = (\mathbf{x}_{Contact}(\mathbf{q}, \mathbf{p}) - \mathbf{x}_{Contact}(\mathbf{q}_{Before}, \mathbf{p}_{Before}))^2. \quad (6.10)$$

The computation of the contact point is actuator specific - it depends on the degrees of freedom supplied. Therefore, the specific implementation of this function will be discussed in the sections specialised on each actuator.

6.5 Modelling for Height-Adjustable Workplaces

This section adapts the general optimisation problem presented in Sec. 6.3 for a height-adjustable platform carrying the workpiece. An example of such an actuator type is an actuated table. The actuator parameters, \mathbf{p} , constitute only one value representing the height of the platform. This height value is expressed by the height difference from the

platform height before optimisation. Also, the term "offset" will be used for this difference. Using this relative height notation rather than the absolute height value facilitates simplifying the terms considering the actuator parameters. The starting point is the general objective function:

$$\begin{aligned}
 f(\mathbf{q}, \mathbf{p}) = & c_1 f_{Posture}(\mathbf{q}) \\
 & + c_2 f_{Centroid}(\mathbf{q}) \\
 & + c_3 f_{Dir}(\mathbf{q}, \mathbf{p}) \\
 & + c_4 f_{Contact}(\mathbf{q}, \mathbf{p}) \text{ with } \sum_i c_i = 1
 \end{aligned} \tag{6.11}$$

First of all, the working direction objective, $f_{Dir}(\mathbf{q}, \mathbf{p})$, is simplified. As there is no degree of freedom to tilt the work piece, the workpiece orientation always remains the same ($\phi_{Dir}(\mathbf{p}) = \phi_{Dir}(\mathbf{p}_{Before})$). This leads to the working direction objective:

$$\begin{aligned}
 f_{Dir}(\mathbf{q}, \mathbf{p}) = & ((\phi(\mathbf{d}_w) - \phi_{Dir}(\mathbf{p})) - (\phi(\mathbf{d}_{wBefore}) - \phi_{Dir}(\mathbf{p}_{Before})))^2 \\
 = & ((\phi(\mathbf{d}_w) - \phi(\mathbf{d}_{wBefore})))^2 \\
 = & (\sum_j \mathbf{q}_j - \sum_j \mathbf{q}_{Before,j})^2.
 \end{aligned} \tag{6.12}$$

Second, the contact point objective, $f_{Contact}(\mathbf{q}, \mathbf{p})$, is simplified. Assuming that the workpiece 2D shape is modelled as a rectangle in perpendicular orientation to the global x- and y-axes, a vertical motion of the actuator does not change the x-direction of the contact point. Therefore, the horizontal component of the contact point can be omitted:

$$f_{Contact}(\mathbf{q}, \mathbf{p}) = (\mathbf{x}_{Contact,y}(\mathbf{q}, \mathbf{p}) - \mathbf{x}_{Contact,y}(\mathbf{q}_{Before}, \mathbf{p}_{Before}))^2 \tag{6.13}$$

Furthermore, the contact point was not expressed in global coordinates, but as the point relative to the bottom of the workpiece or the platform. As mentioned before, the platform height, \mathbf{p} , was modelled as relative height offset by the platform height before optimisation via $\mathbf{x}_{Platform,y}(\mathbf{p}_{Before})$. Combining these two expressions yields:

$$\begin{aligned}
 \mathbf{x}_{Contact,y}(\mathbf{q}, \mathbf{p}) = & \mathbf{x}_{Hand,y}(\mathbf{q}) - \mathbf{x}_{Platform,y}(\mathbf{p}) \\
 = & \mathbf{x}_{Hand,y}(\mathbf{q}) - (\mathbf{x}_{Platform,y}(\mathbf{p}_{Before}) + \mathbf{p}).
 \end{aligned} \tag{6.14}$$

Finally, the height offset was computed by subtracting the height of the hands before optimisation from height after optimisation:

$$\mathbf{p} = \mathbf{x}_{Hand,y}(\mathbf{q}) - \mathbf{x}_{Hand,y}(\mathbf{q}_{Before}). \tag{6.15}$$

Substituting equations 6.14 and 6.15 into 6.13 leads to the result that $f_{Contact}$ is always zero, meaning that the term can be omitted:

$$\begin{aligned} f_{Contact}(\mathbf{q}, \mathbf{p}) &= \left[\left(\mathbf{x}_{Hand,y}(\mathbf{q}) - \left(\mathbf{x}_{Platform,y}(\mathbf{p}_{Before}) + \mathbf{x}_{Hand,y}(\mathbf{q}) - \mathbf{x}_{Hand,y}(\mathbf{q}_{Before}) \right) \right) \right. \\ &\quad \left. - \left(\mathbf{x}_{Hand,y}(\mathbf{q}_{Before}) - \mathbf{x}_{Platform,y}(\mathbf{p}_{Before}) \right) \right]^2 \\ &= 0. \end{aligned} \quad (6.16)$$

As can be seen after simplification, f_{Dir} and $f_{Contact}$ did not depend on \mathbf{p} . Therefore, the posture optimisation task can be solved by strictly optimising \mathbf{q} and inserting the solution, \mathbf{q}^* , into 6.15 to obtain the actuator parameters. Having derived the objectives, the constraints were then designed. To start, the range of the joint angles, \mathbf{q} , was limited to exclude the space of infeasible postures:

$$\mathbf{q}_{min} \leq \mathbf{q} \leq \mathbf{q}_{max} \quad (6.17)$$

For the height-adjustable workplace, the range of height of the actuator needed to be restricted, and this could be realised by adding the inequality constraints

$$h_{min} \leq \mathbf{x}_{Platform}(\mathbf{p}) + \mathbf{x}_{Hand}(\mathbf{q}) - \mathbf{x}_{Hand}(\mathbf{q}_{Before}) \leq h_{max}. \quad (6.18)$$

The final optimisation problem for the height-adjustable platform is summarised by:

$$\begin{aligned} \mathbf{q}^* &= \arg \min_{\mathbf{q}} \|\mathbf{q} - \mathbf{q}_{ideal}\|^2 \\ &\quad + (\mathbf{x}_{Foot,x} - \mathbf{x}_{Centroid,x})^2 \\ &\quad + \left(\sum_j \mathbf{q} - \sum_j \mathbf{q}_{Before,j} \right)^2 \end{aligned} \quad (6.19)$$

subject to

$$\begin{aligned} \mathbf{q}_{min} &\leq \mathbf{q} \leq \mathbf{q}_{max} \\ h_{min} &\leq \mathbf{x}_{Platform}(\mathbf{p}) + \mathbf{x}_{Hand}(\mathbf{q}) - \mathbf{x}_{Hand}(\mathbf{q}_{Before}) \leq h_{max}. \end{aligned}$$

Having obtained \mathbf{q}^* , the height offset, \mathbf{p}^* , can be found with:

$$\mathbf{p}^* = \mathbf{x}_{Hand,y}(\mathbf{q}^*) - \mathbf{x}_{Hand,y}(\mathbf{q}_{Before}) \quad (6.20)$$

6.6 Modelling for Robot Kinematics

In this section, the objective function is adapted to solve the posture optimisation problem for six-axis robot actuators. Among the six axes, only three are used. As with the human model, the robot actuator is described as a kinematic chain containing of segments and parametrised with a joint angle vector, \mathbf{p} . Determining the robot parameters required an

inverse kinematics computation. In order to directly obtain \mathbf{p} , it was integrated into the optimisation problem.

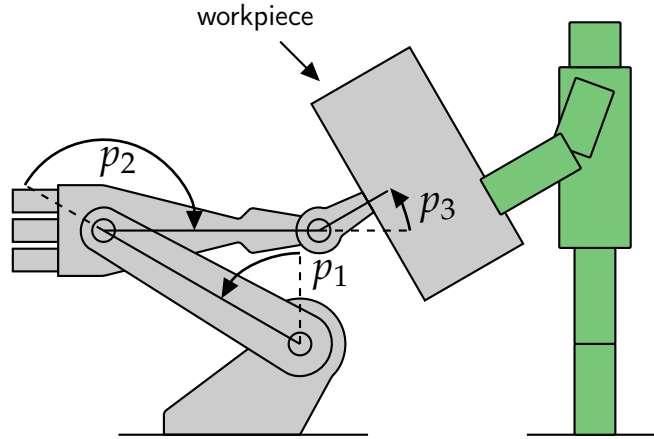


Figure 6.8: Robot model notation used to describe the proposed approach.

Again, the modelling started with the formulation of the general optimisation problem:

$$\begin{aligned} f(\mathbf{q}, \mathbf{p}) = & c_1 f_{Posture}(\mathbf{q}) \\ & + c_2 f_{Centroid}(\mathbf{q}) \\ & + c_3 f_{Dir}(\mathbf{q}, \mathbf{p}) \\ & + c_4 f_{Contact}(\mathbf{q}, \mathbf{p}) \text{ with } \sum_i c_i = 1. \end{aligned}$$

In contrast to the height-adjustable platform, the simplification of the work direction objective could not be applied here. Hence, the full definition stated in Eq. 6.9 had to be used. The calculation of the contact point had to consider both human and actuator joint angles. Let \mathbf{x}_{Eff} be the position and \mathbf{d}_{Eff} the orientation of the end effector. Moreover, let w_{Dir} represent the width of the workpiece. The object front point, \mathbf{x}_{Obj} (see Fig. 6.9), is then introduced. It is computed as:

$$\begin{aligned} \mathbf{x}_{Obj} &= \mathbf{x}_{Eff} + w_{Dir} \phi(\mathbf{d}_{Eff}) \\ &= \mathbf{x}_{Root} + \sum_{j=1}^{J_{Robot}} \mathbf{R} \left(\sum_{i=0}^j \mathbf{p}_i \right) * h_{Robot,j} + w_{Dir} \sum_j^{J_{Robot}} \mathbf{p}_j. \end{aligned} \quad (6.21)$$

Further the direction vector perpendicular to the end effector orientation, $\mathbf{d}_{Eff\perp}$ was computed:

$$\mathbf{d}_{Eff\perp} = \frac{1}{\|\mathbf{d}_{Eff}\|} \cdot \begin{pmatrix} \mathbf{d}_{Eff,y} \\ -\mathbf{d}_{Eff,x} \end{pmatrix} \quad (6.22)$$

The contact point was defined as the point where the human hand and the workpiece touch. Especially interesting for the optimisation was the y component, or the height of the contact point. The x -component can always be adjusted as the human being can move horizontally. This contact height was encoded relative to the object front, \mathbf{x}_{Obj} , because the

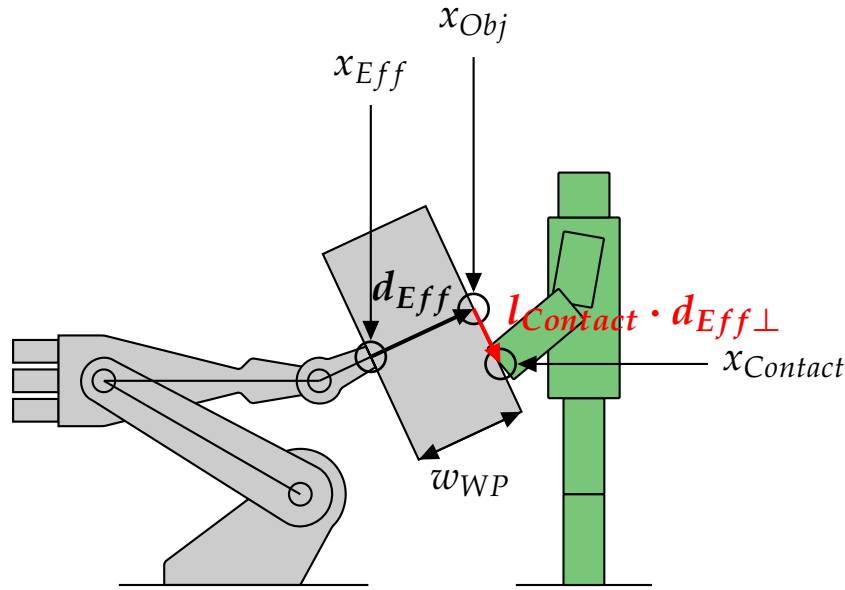


Figure 6.9: Visualisation of the variables used in the optimisation problem for the robot actuator.

region of the workpiece to be processed was of interest and not the absolute hand height. Hence, the object front was allowed to move, however, the contact point had to remain statically positioned. The condition described before can be expressed by the equation:

$$\begin{aligned} \mathbf{x}_{Hand,y}(\mathbf{q}) &= \mathbf{x}_{Obj,y}(\mathbf{p}) + \mathbf{x}_{Contact,y}(\mathbf{q}, \mathbf{p}) \\ &= \mathbf{x}_{Obj,y}(\mathbf{p}) + l_{Contact}(\mathbf{q}, \mathbf{p}) \cdot \mathbf{d}_{Eff\perp,y}(\mathbf{p}) \end{aligned} \quad (6.23)$$

where $l_{Contact}$ denotes the (signed) length from \mathbf{x}_{Obj} to $\mathbf{x}_{Contact}$ along the object front (see Fig. 6.9 marked in red). Isolating $l_{Contact}$ in Eq. 6.23, yields:

$$l_{Contact}(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{x}_{Hand,y}(\mathbf{q}) - \mathbf{x}_{Obj,y}(\mathbf{p})}{\mathbf{d}_{Eff\perp,y}(\mathbf{p})}. \quad (6.24)$$

Equation 6.24 was then used in the contact point objective Eq. 6.10:

$$f_{Contact}(\mathbf{q}, \mathbf{p}) = (l_{Contact}(\mathbf{q}, \mathbf{p}) - l_{Contact}(\mathbf{q}_{Before}, \mathbf{p}_{Before}))^2$$

In short, the contact point objective for the robot suggested that the signed height difference of the contact point relative to the object front point (see Fig. 6.9 red arrow) was to remain similar. Note that this condition assumed that the worker processes the object front, though if the worker processes a different side of the workpiece, further object front points could be introduced for the top and bottom sides. Once the objective function was derived, the constraints were designed. At first, there were minimum and maximum

angles for human and robot joints.

$$\begin{aligned} \mathbf{q}_{min} &\leq \mathbf{q} \leq \mathbf{q}_{max} \\ \mathbf{p}_{min} &\leq \mathbf{p} \leq \mathbf{p}_{max} \end{aligned} \quad (6.25)$$

When tilting the workpiece, a problem could surface - posture and contact point may be correct, however, the posture would still not be feasible as parts of the body collide with the workpiece. Hence a collision constraint had to be introduced. The reason for this constraint is that every point of the body has to reside on the side of the object front different from the robot front. This can be carried out using the Hessian normal form of the object front. The constraints ensure that each point of the body (e.g., the front of the head) substituted in the object front plane definition is always positive.

$$(\mathbf{x}_{Head}(\mathbf{q}) - \mathbf{x}_{Obj}(\mathbf{p})) \cdot \mathbf{d}_{Eff\perp} \geq 0. \quad (6.26)$$

Briefly, the optimisation problem for the robot kinematics is defined as:

$$\begin{aligned} \mathbf{q}^*, \mathbf{p}^* = \arg \min_{\mathbf{q}, \mathbf{p}} & (\mathbf{q} - \mathbf{q}^{Before})^2 \\ & + (\mathbf{x}_{Foot} - \mathbf{x}_{Centroid}(\mathbf{q}))^2 \\ & + ((\sum_j^J \mathbf{q}_j - \sum_j^{J_{Robot}} \mathbf{p}_j) - (\sum_j^J \mathbf{q}_{Before,j} - \sum_j^{J_{Robot}} \mathbf{p}_{Before,j}))^2 \\ & + (\mathbf{x}_{Contact,y}(\mathbf{q}, \mathbf{p}) - \mathbf{x}_{Contact,y}(\mathbf{q}_{Before}, \mathbf{p}_{Before}))^2 \end{aligned} \quad (6.27)$$

subject to

$$\begin{aligned} \mathbf{q}_{min} &\leq \mathbf{q} \leq \mathbf{q}_{max} \\ \mathbf{p}_{min} &\leq \mathbf{p} \leq \mathbf{p}_{max} \\ (\mathbf{x}_{Head}(\mathbf{q}) - \mathbf{x}_{Obj}(\mathbf{p})) \cdot \mathbf{d}_{Eff\perp} &\geq 0. \end{aligned}$$

6.7 Solving the Optimisation Problem

Having defined the optimisation problems, an algorithm was required to obtain the solutions. The proposed formulations were an optimisation problem with non-linear constraints. Specifically with regards to the computation of the position vectors, \mathbf{x}_{Head} and \mathbf{x}_{Hand} , they needed the application of a non-linear forward kinematics function in the angular space. Based on the fact that objective functions and constraints contained trigonometric functions from the rotation matrix, the stated problems were also non-convex and including multiple local minima. The choice of solvers for non-linear problems with non-linear constraints is limited to a subset of algorithms. Algorithms implementing the Sequential Quadratic Programming (SQP) approach (see Sec. C of the Appendix for a detailed explanation) were utilised. SQP is widely used for non-linear constrained optimisation problems, especially in the field of posture prediction [8, 80]. A superb report of the

algorithm framework can be found in numerical optimisation literature, such as Nocedal and Wright [193]. The main concept is to iteratively find the solution by successively approximating the original non-linear program with a quadratic program containing linear constraints.

Considering initialisation, various postures were tested, including the ideal posture, \mathbf{q}_{ideal} , the current posture, \mathbf{q} , and convex combinations thereof. Empirical analyses indicated that initialising with the ideal posture, \mathbf{q}_{ideal} , generated the most favourable results for the present task.

6.8 Control Strategies

In this section, the design of a so-called control strategy, which decides whether to propose a workplace layout to the worker or not, is examined. Its purpose is to keep the postural load as well as the amount of workflow interruptions emanating from workplace adjustments at an acceptable level. As stated in Secs. 6.2.3, the control strategy outputs the actuator configuration, \mathbf{p} , at each point of time, t . As such, there are two possible cases - either the strategy decides to propose a layout change by activating an alternative actuator configuration or it settles upon maintaining the layout. The derivation of the alternative actuator configuration was already discussed in Sec. 6.5 and 6.6. Here, the design of the algorithm is explained. In particular, an answer is provided addressing the question of how the method decides when to activate the configuration and when not to. The control strategy was integrated into the WPC loop (see Fig. 6.12).

The task was to design a decision function, $d_{adj}(P(t)) \in 0, 1$. This function d_{adj} returns 1, to ascertain when the WPC should notify the worker with proposition of adaptation of the workplace and 0 when the current layout would best be maintained. The decision was based on the current posture sequence, $P(t) = \{P_1, \dots, P_t\}$, obtained from the posture classification component. The aim was to keep the final EAWS score, $S_{EAWS}(t_{End})$, as well as the accumulated distance of the actuator path (ADAP) at a reasonable level. The ADAP is introduced as a measure of how much the actuator moved. The measure can be directly transformed into the interruption time. The ADAP denotes the length of the path the actuator travelled. For example, for a height-adjustable actuator, the ADAP is the sum of all height offsets during the WPC loop. For a robot actuator, one could use the length of the path the end effector travels, including its rotation. Alternatively, the angular displacements of the robot joints could be accumulated. The challenge was to predict which type of effect an alternative configuration would have on the final EAWS score. Omitting the layout adjustment can have a fatal impact on the ergonomic assessment as well as be negligible. In the following, two decision strategies are proposed which attempt to predict the effect of the current posture sequence $P(t)$ on the final EAWS score and employ them to build the decision function, d_{adj} .

6.8.1 Temporal EAWS Threshold (TET) Control Strategy

The EAWS computes the score at the end of the recording by analysing the time share of each posture class adopted. The principal rationale of this strategy was to use the temporal EAWS score, $S_{EAWS}(t)$, to derive a decision. The temporal EAWS score is computed by applying the EAWS computation scheme to the posture sequence adopted from the beginning of the recording until the current point in time, t . When $S_{EAWS}(t)$ exceeds an acceptable threshold, T , the strategy proposes the alternative actuator configuration:

$$d_{adj}(P) = \begin{cases} 1, & \text{if } S_{EAWS}(t) > T, \\ 0, & \text{else.} \end{cases} \quad (6.28)$$

$S_{EAWS}(t)$ at the current time, t , is computed by applying the EAWS computation scheme (see Fig. A.2 of the Appendix) on the current posture sequence, $P(t)$. The lower the value of T , the more likely the final EAWS score is reduced. Figure 6.10 illustrates the relationship between T and S_{EAWS} using an exemplary simulated process (see Sec. 8.4.2 for realisation of the simulation process). A worker periodically adopts a sequence of postures to accomplish the tasks at hand. The workpiece is assumed to be height-adjustable. The progression of the EAWS score is depicted by the blue line whereas the red line indicates the movement of the actuator.

Figure 6.10a represents the case where no adjustment is initiated. After around 400 seconds, a pattern of the EAWS curve can be identified indicating that the worker periodically alternates between ergonomically different postures. However, the EAWS score always remains at an unacceptable level (above 50 points).

Figure 6.10b shows the case where an adjustment is initiated every time the EAWS score exceeds the value 30. As can be seen, the score remains below 50 points as the working height is continuously adapted to match the current task. However, the number of adjustments increases, leading to a higher ADAP. Furthermore, the figure suggests that the score changes less the longer the process endures. This behaviour is typical for the EAWS (see Sec 6.8.2 for explanation). The implication is that if the temporal score settles to a value above the given threshold, as in the present case, the control strategy will continuously suggest an adjustment. In conclusion, ineffectual workplace adaptations will be made in the later stages of the process.

To sum up, the implicit assumption of the Temporal EAWS Threshold (TET) strategy is that the EAWS score at the end will be basically the same as the current one. When there is no interference from the WPC, the ergonomic assessment result will remain.

6.8.2 Decaying Temporal EAWS Threshold (DTET) Control Strategy

One property of the EAWS calculation scheme is that the longer the recording endures, the less the final score changes (see Fig. 6.10b). A simple example which explains how

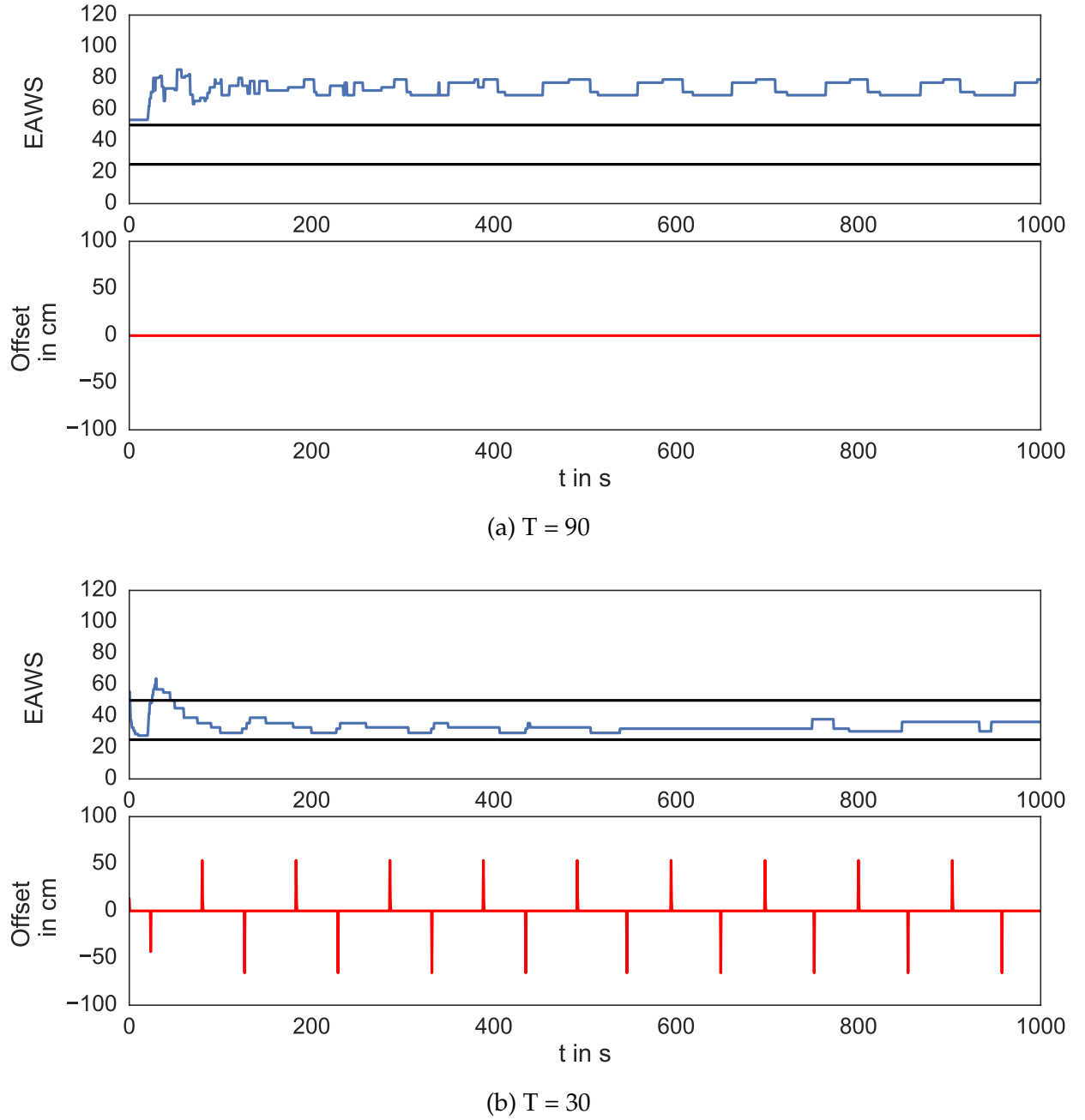


Figure 6.10: Visualisation of EAWS score (black lines denote EAWS scores 25 and 50) and height-adjustable actuator offset over time for the TET strategy.

to calculate the EAWS has been presented in Fig. A.2 of the Appendix. The final EAWS score is computed as the sum of partial EAWS scores for each posture class, p :

$$EAWS = \sum_p EAWS_p. \quad (6.29)$$

Each partial score, $EAWS_p$, is dependent on the time proportion, τ_p , spent in posture, p . The partial score is computed through a lookup-table function, LUT (see Appendix A.1 for the complete EAWS lookup table), which assigns a partial score depending on the bin

τ_p is located in:

$$EAWS_p = LUT_p(\tau_p). \quad (6.30)$$

Given the entire process as a sequence of n image frames and the posture labelling for each, the time proportion, τ_p , can be computed as:

$$\tau_p = \frac{n_p}{n}. \quad (6.31)$$

where n_p refers to the number of frames containing posture, p . When the worker adopts a new posture for Δn_p frames, the fractional change, $\Delta\tau_p$, can be computed as:

$$\Delta\tau_p = \frac{n_p + \Delta n_p}{n + \Delta n_p} - \frac{n_p}{n}. \quad (6.32)$$

The longer the recording time, n , the less influence Δn_p has on $\Delta\tau_p$. The Decaying Temporal EAWS Threshold (DTET) strategy begins with a rather low temporal EAWS threshold, T_{Start} , and increases it over time until a final threshold, T_{End} . The current threshold, T_t , at time, t , is computed as:

$$T_t = \begin{cases} T = T_{Start} + \frac{t}{t_{Total}} (T_{End} - T_{Start}) & , \text{ if } t < t_{Total}, \\ T_{End}, & , \text{ else.} \end{cases} \quad (6.33)$$

The parameter, t_{Total} , indicates how fast the threshold approaches the final value. The computed threshold, T_t , was utilised as a threshold in the TET control strategy explained in Sec. 6.8.1. The longer the process endures, the less the WPC intervenes - the final EAWS is less influenced. To illustrate this effect, Fig. 6.11 depicts the progression of EAWS score and actuator movement using the DTET strategy. The underlying process is the same as used in Fig. 6.10. It is worth pointing out that the frequency of adjustments gradually decays until actuator movements completely cease after roughly 700 seconds. In exchange, the EAWS score first decreases, but then continues to grow due to a lack of intervention.

A further modification was added to the DTET control strategy. As seen in Figure 6.11, the actuator activity stopped towards the end of the recording. The last used actuator configuration was then fixed for the rest of the process. This configuration could have been unsuitable for the remaining postures to be adopted. In assuming that the remaining posture sequence was similar to the former, a standard actuator configuration was initiated after t_{Total} was reached. As standard configuration, the mean or the median of all actuator configurations set until the current time, was employed.

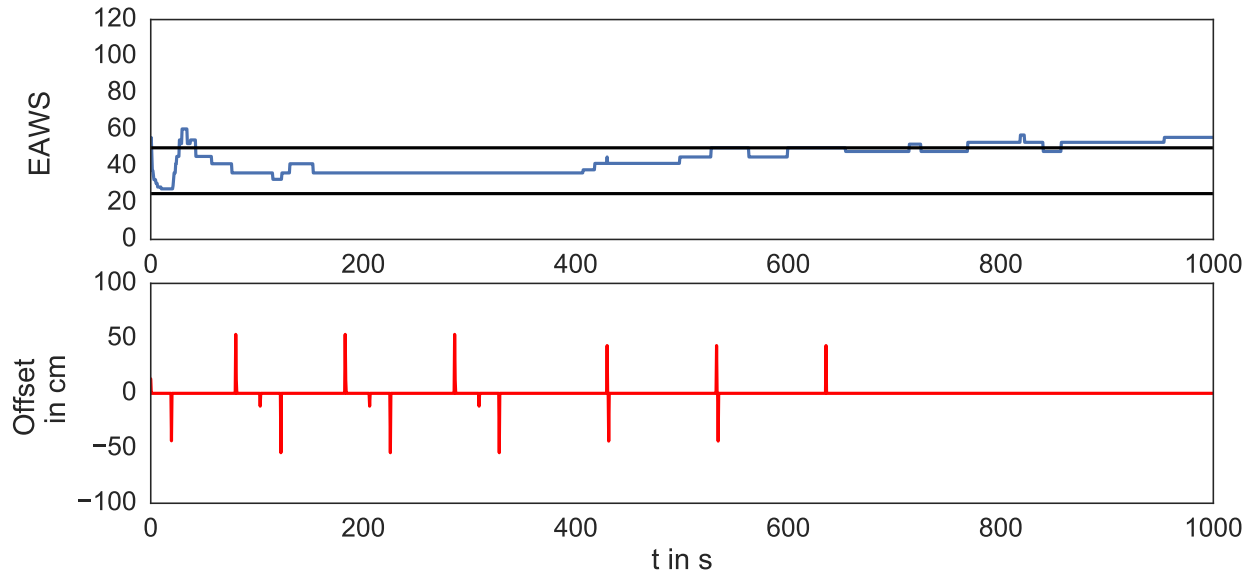


Figure 6.11: Visualisation of EAWS and height-adjustable actuator offset over time for the DTET strategy.

6.9 Conclusions

This chapter has covered how to compute workpiece pose configurations, enabling comfortable working conditions. The major challenge was to predict the worker's posture after adjustment.

The optimisation-based approach was elected as it seamlessly integrates into the workpiece pose search task. Similarly, the workpiece pose search task was solved by stating an optimisation problem with a solution that represented the actuator configuration. In addition to the actuator configuration, the proposed posture was computed. To find the appropriate workpiece posture, the application of the SQP algorithm was suggested. The proposed approach was adapted for a height-adjustable actuator along with a robot actuator. In effect, both algorithms differed with regards to implementation of the working direction objective and the contact point objective.

Furthermore, this chapter has put forward that the dynamic behaviour of the WPC has to be accounted for. Frequent adjustments unacceptably slow down the workflow. It is important to optimise ergonomics as well as minimise actuator activity through a chosen control strategy. This problem can be interpreted as an optimal control problem, however the approaches seen in the literature are not suited for the type of dynamic system this thesis is focused on. Hence, the TET and DTET control strategies were proposed. Both exploit the properties of the EAWS method in order to reduce the score while keeping the interruptions at a reasonable level.

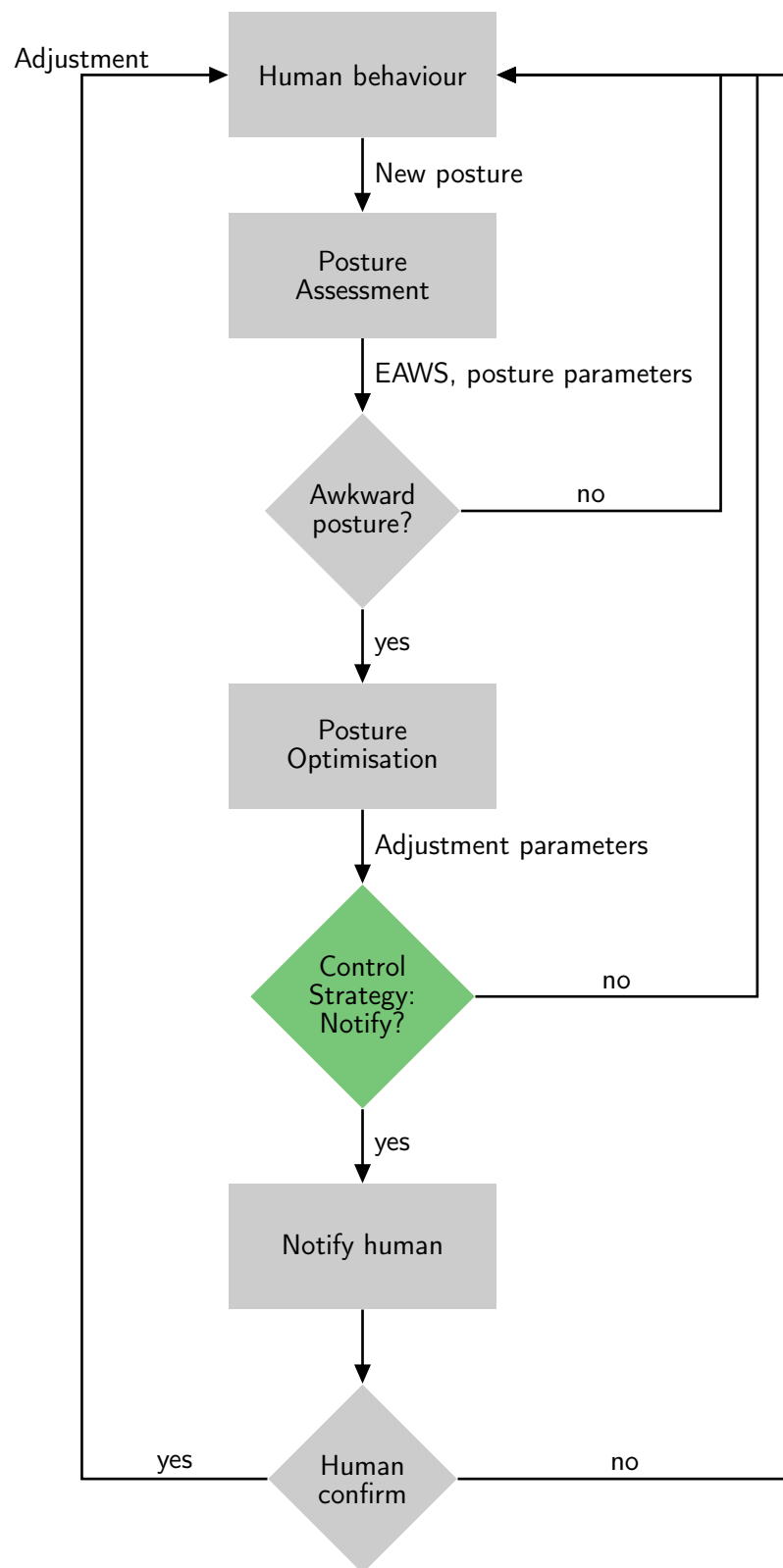


Figure 6.12: Control strategy integrated into the WPC loop.

7 Implementation

This chapter describes the implementation details of the WPC. Specifically, the software architecture used to build the WPC is explained. Besides this, the chapter outlines how the methods developed were integrated into a demonstrator scenario.

7.1 Scenario

Consider an assembly task where a worker is processing a workpiece. The workpiece is fixed on an actuator that can manipulate its spatial pose. The worker performs tasks on different regions of the workpiece. The workpiece dimensions should be large enough that it is not possible to perform all tasks in an ergonomic posture without any workpiece adjustment. An example could be a bicycle frame as depicted in Fig. 7.1. If tasks must be performed at the top of the frame as well as at its bottom, the user normally needs to bend forward as well as raise their arms to shoulder height. As soon as the worker enters the sensors' FOV, the background subtraction algorithm (see Sec. 5.5.1) segments the worker. A foot switch is placed under the workpiece and when the user presses the button, the joint localisation (see Sec. 5.5.3) is initialised. During this initialisation phase, body sizes are estimated from the worker's height. Hence, the worker is expected to stand upright when activating the foot switch. After initialisation, the WPC loop starts.

Once the WPC loop has begun, the worker performs the tasks at hand while the system records his or her posture. The system displays the RGB image stream with an overlay of the worker's posture to provide visual feedback to the user. The WPC notifies the worker when the control strategy proposes an adjustment of the workpiece pose. The worker can either confirm or decline the adjustment. Upon agreement, the system waits until the worker steps out of the scene and pushes a confirmation button to initiate the adjustment. In the case of a manually adjusted actuator, the WPC overlays markers on the screen in order to guide the worker in adjusting the workplace. After adjustment, the WPC loop starts anew.



Figure 7.1: The WPC demonstrator scenario.

7.2 Hardware

This section describes the hardware devices necessary to build the demonstrator scenario in Sec. 7.1. First, a RGBD sensor is needed to capture the scene. A Microsoft Kinect® sensor v1 was chosen, which is simply referred to as Kinect in this work. When research in the scope of this thesis started, the Kinect v2 or other notable RGBD devices were not yet available. A computer is needed to process the image stream and compute and initiate adjustments. The WPC algorithms were tested on a laptop containing a two-core i5 CPU (Intel Corp., Santa Clara, CA, USA). For the 3D localisation, a desktop computer containing an i7 CPU (Intel Corp., Santa Clara, CA, USA) and a Geforce GTX 580 GPU (Nvidia Corp., Santa Clara, CA, USA) were used. A foot switch was employed as a control interface for the worker.

Finally, the WPC system needed actuators to adjust the workpiece pose. The actuators tested were a bicycle frame holder where the worker could manually tilt the workpiece along the z-axis (see Fig. 7.2) and a lightweight robot manipulator (see Fig. 7.3). In the following, the components are explained in more detail.

7.2.1 Sensor and Input Device

Kinect v1 Sensor

The Kinect v1 is a low-cost RGBD sensor that was initially released in 2010. Originally, it was utilised to accommodate gesture input for video games. The bundled software offers marker-less joint localisation from the front perspective, which has made the device interesting for human-machine interaction applications beyond video games. The sensor acquires colour images with a resolution of 640x480 pixels and depth images of 320x240

pixels via the structured light approach (see Sec. 5.2.3). The average frame rate is around 30 frames per second. The sensor was placed on a tripod at approximately 30cm height with about 3.5m of distance from the worker.

Foot Switch

A USB Triple Foot Switch II (Scythe Co., Ltd., Tokyo, Japan) provided the user an input interface where the hands would remain free. The foot switch provides three programmable buttons that emulate a mouse or keyboard input. The supplied software allows assigning each button with a keyboard or mouse button. This has the advantage that when debugging the program in the laboratory, foot switch activations could be simulated by pressing the keyboard.

7.2.2 Actuators

Bicycle Frame Holder

The bicycle frame holder was used to fix a bicycle frame such that it stood in front of the worker. The frame can be tilted by manually releasing the lock and rotating the frame. The workpiece pose can be set from almost horizontal to vertical, which yields a tilting range of almost 90° (see Fig 7.2 for visualisation of the range). Considering common bicycle frame dimensions, the lower part of the frame can be raised up to around 30cm.

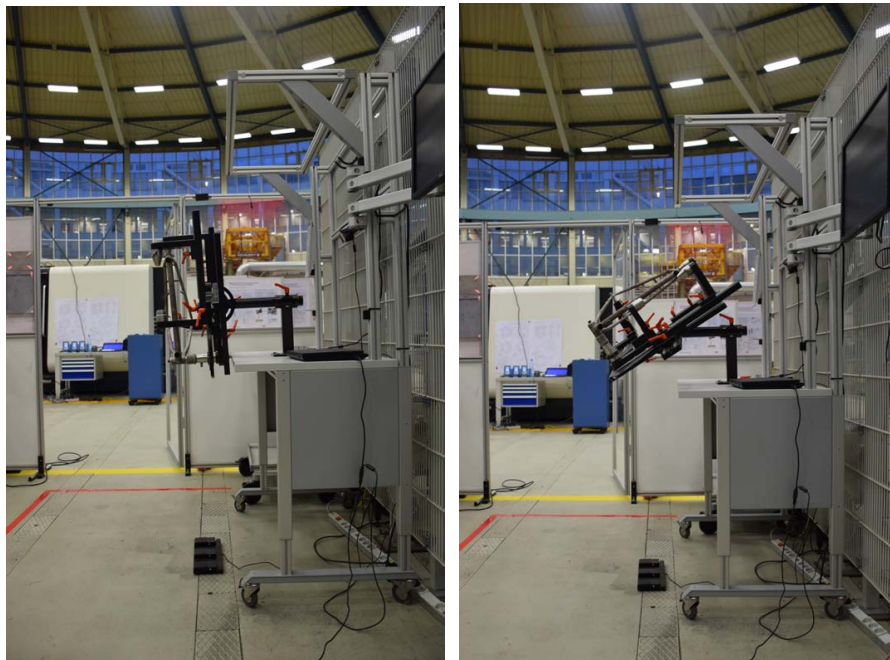


Figure 7.2: Motion range of the bicycle frame holder.

Light-Weight Robot Manipulator

To demonstrate the WPC function with a robot manipulator, the Universal Robot UR5 manipulator [194] was utilised. The UR5 has a payload of 5kgs and a maximal reach radius of 850mm. As the maximum payload is stated to be 5kg, including gripper and sensors, a real workpiece could not be used, but instead only a dummy object.



Figure 7.3: Robot actuator used in the demonstrator scenario.

7.3 Software

This section outlines the software architecture of the WPC and reports the concepts behind. Section D.0.1 of the Appendix describes the concrete tools and implementations used.

7.3.1 WPC System Architecture

The C++ language dominates the WPC code. C++ offers object oriented programming features that permit building reusable modules. As the high-level code is compiled into optimised low-level instructions, the execution performance is faster in comparison to interpreted languages. One limitation is that low-level instructions mostly cause operating system dependence. However, libraries with an abstraction layer help to maintain the code platform independence.

The software architecture was designed with three goals in mind. For one, the code has to be easily understandable and result in a reduced learning effort for a developer reusing or extending the software. As well, modules should be readily integrated into other projects and platforms, and this includes an easy deployment of the software to other computers. Finally, if a user needs a new feature, the code must offer aids to extend it. In summary, the goals were:

- clarity,
- re-usability,
- extensibility.

This is why the code was structured and modularised into isolated units named “packages” (see Sec. D.0.2 of the Appendix for details). The WPC system software is divided into four packages:

- RGBDSensor
- PostureLib
- WPCApp
- Experiments

Figure 7.4 provides a visual overview of the dependent packages.

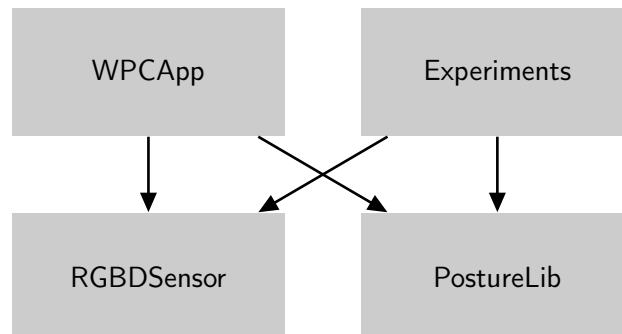


Figure 7.4: Package dependencies within the WPC.

The RGBDSensor package contains classes to interface RGBD sensors to obtain colour and depth images. Currently, an interface to the Kinect v1 is implemented that wraps the libraries, OpenNI2 and NiTE2. The package also features routines to obtain human segmentation masks and joint coordinates with the NiTE2 middleware. Furthermore, the package can emulate a camera by successively reading images from an input folder. This function supports creating experimental setups to validate the algorithms developed.

The PostureLib package contains implementations of algorithms to assess and optimise postures. It offers data structures representing the human kinematic chain for 2D joint localisation and the kinematic tree for 3D joint localisation. The posture optimisation

classes interface the numerical optimisation library, NLOpt. The algorithms have been sped up using parallel processing techniques (see Sec. D.0.3 for details).

The Experiments package has tools that allowed to conduct experiments on the WPC system (see Chapter 8). Specifically, it contains executables to acquire joint coordinates from the developed methods as well as those from the NiTE2 library. Furthermore, the temporal EAWS score and the optimised posture can be recorded for each frame of an image stream. The tools can process multiple datasets in a row with different algorithm parameters. Configurations are read from xml files to avoid re-compiling the code in case the parameters change. Furthermore, the tools can take snapshots from visualisations from every n-th frame to provide data for qualitative analysis of the results. The experiment tools are normally applied to recorded data rather than live data in order to guarantee reproducibility.

The package, WPCApp, features an executable that integrates the algorithmic blocks into the WPC loop. It depends on the packages RGBDSensor and PostureLib. A Graphical User Interface (GUI) visualises the current posture, optimised posture and current EAWS score overlayed onto the current camera image. The WPCApp also offers functions to record image data.

7.3.2 The WPCApp

The WPCApp integrates posture assessment, optimisation and control strategy into one system. This executable file controls the WPC hardware. One of the goals for the WPC system was to be modular and capable of being applied to different actuators, posture assessment algorithms and control strategies.

The WPCApp GUI

The WPCApp GUI (see fig. 7.5 for a screenshot) provides the user with visual feedback. It displays the current RGB image overlayed with the current EAWS score, the current posture, the optimised posture and the proposed actuator adjustment. Hence, it shows the correct posture for the task at hand and explains to the user how the environment will change.

The Model-View-Controller (MVC) Design Pattern

The WPC was designed as a generic system which can be used with various input and output components. Inputs can origin from different types of cameras or user input devices. Physical actuators, their virtual visualisations and the GUI are output components. Whereas the core behaviour of the WPC remains the same, input and output components vary among different configurations. In order to foster flexibility enabling different input and output components to be plugged into the system without changing the core, the Model-View-Controller MVC design pattern [195] was employed. The MVC is an archi-

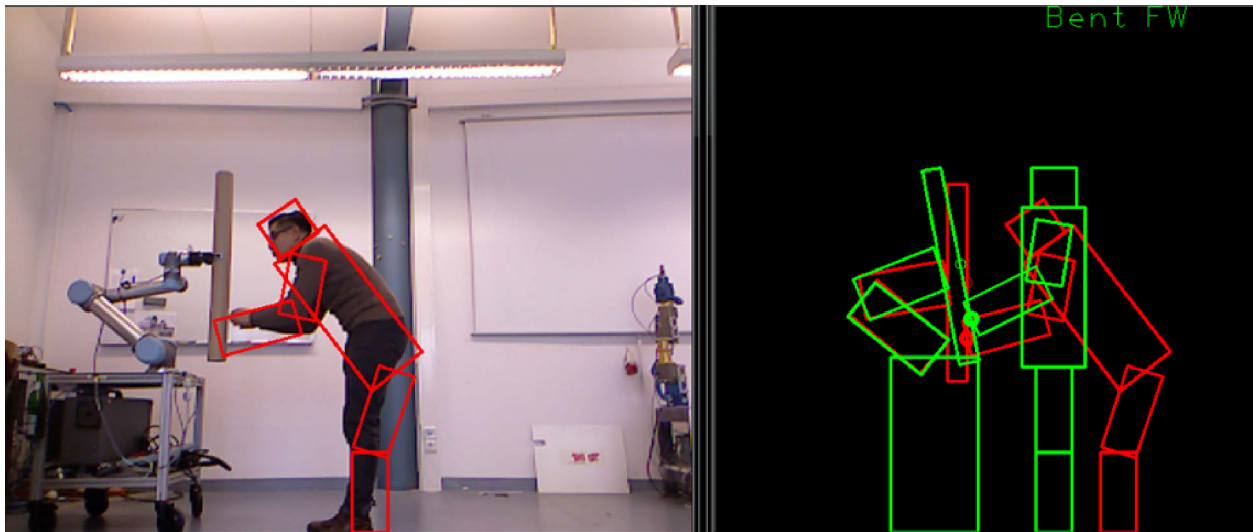


Figure 7.5: The WPCApp graphical user interface.

tectural design pattern mostly found in user interfaces. It separates the internal representation of the data and program logic from the user interface. The MVC pattern proposes dividing the system into three interconnected components: model, view and controller.

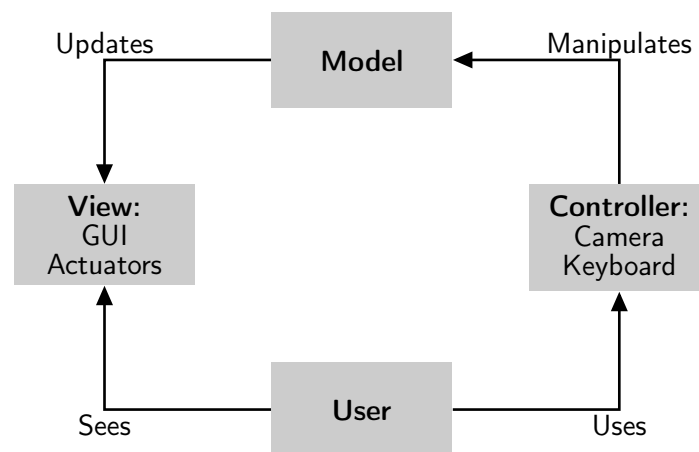


Figure 7.6: Visualisation of the MVC pattern.

The model assimilates the state of the system and the data stored. The state of the WPC system is represented by variables, such as the current EAWS score, posture sequence or actuator configuration. Furthermore, the model possesses program logic, like posture assessment and optimisation routines, that manipulate the state of the WPC system. Multiple controllers handle the inputs from the user or devices. Inputs can be new frames from the sensor as well as user input. Exemplary controller objects are the camera or the keyboard input. Finally, there can be one or multiple views. Each view presents particular parts of the model state to the user. When the model has been changed, it updates the views. Examples for the views are GUI visualisations or the physical actuator control. The communication interface between models, views and controllers is standardised, such that different implementations of views or controllers are interchangeable.

Communication between the MVC Components

Concerning the communication workflow of the MVC components, the focus was put on two requirements: For one, the flow of information was to be managed automatically. Once one model receives user or camera device inputs, it executes the corresponding program routines, which change its state variables. Once particular state variables have been modified, the model updates the views interested in these variables. As there is a large amount of possible control paths to be activated, automatically triggering each component was preferred to having their order of execution hard-coded. Second, over the course of the development process, the structure of the model and controller frequently change. For example, the set of state variables can be extended or there are new parts in the program logic. Hence, it is necessary to achieve the least dependency possible between each component. These two requirements can be met using the observer pattern [196]. The observer pattern separates the interacting components into two groups - subject and observer. Observers register on subjects. A subject can notify all of its observers, of which their "update" methods are then executed. A component is incorporated into the system by deriving from the subject or observer base classes implementing the required interfaces. It is possible that components are subjects as well as observers, as necessary in the case of the model. Figure 7.7 shows an example of how a model and a view interact with each other within the observer pattern.

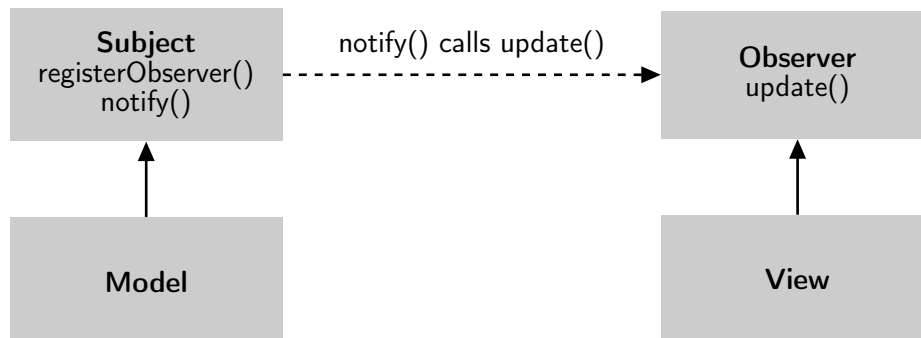


Figure 7.7: Observer pattern applied when the model updates the view.

The model is registered as the observer of the controller and the views are registered as observers of the model. So as to not notify all views at each change in the subject, the views register to specific events rather than the model itself. Events can then be registered to specific parts of the model or control components. Using this event system, the components can vary without breaking the interfaces between one another. Furthermore, the computation overhead is reduced as only the parts to be updated are actually notified.

8 Evaluation

This chapter presents and discusses the results of the experimental evaluations conducted on the components of the WPC system. Before delving into the setup and results, the purpose of the evaluations is stated. The central research goal of this thesis has been delineated in Chapter 3 as:

*It is possible to realise an **in-process** WMSD prevention measure addressing the risk factor, **posture**, at an **intensity level of 1 - 2** on the effectiveness scale according to Goggins et al. by combining **image-based ergonomic assessment** with **in-process workplace layout design**.*

The purpose of the experiments is to determine whether the self-stated goal has been achieved. Thereby, the focus lies on analysing the aspects marked in bold. First, there is a need to prove that the algorithms were able to supply immediate feedback, such that the system can directly support the worker during the process. Secondly, the system shall reliably detect a critical situation by image-based ergonomic assessment. In the best case, the assessment results shall be comparable to what a human being would conclude. Finally, the postural risk level should have been eliminated, or at least reduced, through workplace layout adjustments proposed. An experimental setup which shows that these three aspects apply has to comprise three groups of experiments:

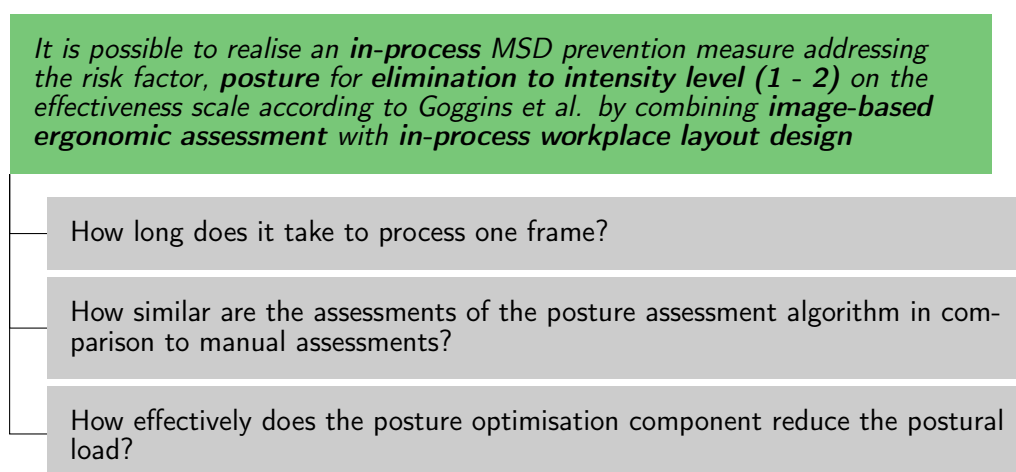


Figure 8.1: Research questions to be answered in the evaluation experiments.

8.1 Experimental Datasets

Having stated the required experiments, the datasets for evaluation were collected. The sensor was placed at a distance of roughly 3.5m from the scene, recording the subject from a side perspective. In total, three datasets were generated that concentrated on the specific evaluation aspects. Table 8.1 summarises the datasets created and the features they address.

Table 8.1: Summary of datasets (number of sequences and content) and their usage in the experiments.

Dataset	Seqs	Content	Purpose
DAT1	9	Image data of standing posture sequences of an approximately one minute duration.	<ul style="list-style-type: none"> • Measure processing time per frame. • Compare posture assessment of standing postures in a sequence with human assessment. • Evaluate standing postures before and after posture optimisation.
DAT2	32	Image data of standing, sitting and kneeling postures in isolation, including motion capture data.	<ul style="list-style-type: none"> • Compare posture assessment of standing, sitting and kneeling EAWS postures with a commercial DHM tool.
DAT3	18	Joint angle sequences of standing postures with a duration of roughly 24 minutes.	<ul style="list-style-type: none"> • Determine the relationship between EAWS and ADAP for different control strategies at a height-adjustable actuator.

DAT1: Posture Sequences

The first dataset (DAT1) was employed to evaluate processing time, posture assessment as well as posture optimisation. Figure 8.2 portrays sample frames from DAT1. The dataset contains image sequences of nine subjects (one female and eight males) with anthropometric heights roughly between 1.5m and 2m performing the five different stand-

ing EAWS postures ("standing", "bent forward", "strongly bent forward", "upright with elbow at shoulder level" and "upright with hands above head level"). The postures adopted in this dataset were intended for assessing the algorithms with reference to the demonstration scenario presented in Sec. 7.1. Thus, either a height-adjustable actuator or a robot actuator were expected to hold the workpiece. As the degrees of freedom were mostly suitable for standing postures, sitting and kneeling postures were excluded from this dataset. Each subject adopted a sequence of postures in a randomised order. The duration of each posture was arbitrarily chosen to amount to between one and three seconds. The hand heights recorded ranged from 0cm (strongly bent) to roughly 250cm (hands above head). The postures from the video sequences were classified by a person different from that who selected the order of the postures. With this design, it was enforced that the human classifier only assigns labels based on the image data and not on the known sequence.

DAT2: Complete Posture Set

The second dataset (DAT2) features more details on the posture assessment. Of particular interest was the question of how accurately the posture was predicted in comparison to current DHM tools using motion capture systems. In contrast to DAT1, DAT2 contained 11 EAWS postures, excluding "standing and walking in alteration", "sitting upright with back support", "sitting upright with hands above head level", "lying on back" and "climbing". The postures were performed by three male subjects with anthropometric heights roughly ranging from 160cm to 180cm. Figure 8.3 shows exemplary frames.

The postures adopted were simultaneously recorded by the Kinect and a Vicon MX3+ motion capture system (Vicon Motion Systems Ltd., Oxford, UK). The Vicon system requires that the subject recorded wears a set of markers located at pre-defined positions of the body (see Fig. 8.4a for marker placement setup). The system then tracks these markers at a spatial resolution of 100Hz, which are then used to automatically reconstruct a full body motion (see Fig. 8.4b). This motion is then classified into postures using the commercial tool Alaska/EAWS [70] (see Fig. 8.4c). A single posture per video sequence was recorded instead of a whole sequence. This ensures that frames where the subject transitions from one posture to another are not considered in the evaluation. In total, 32 videos were obtained where the subjects performed a single posture each.

DAT3: Processes over Several Minutes

The third dataset (DAT3) was utilised to evaluate the control strategies developed for the WPC. As a basis, the posture sequences from DAT1 were used. While being sufficient to evaluate the principal functions of the WPC, they do not realistically simulate a process as the total duration of each sequence of nearly one minute is too short. Sample processes of several minutes were created from the DAT1 posture sequences by repeating them and lowering the number of frames per second. DAT3 does not contain image data - only



(a) RGB images



(b) Silhouettes

Figure 8.2: Sample frames from DAT1.

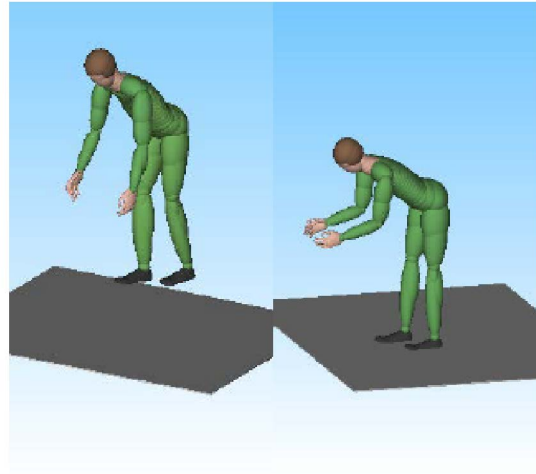
the tracked joint angles. This dataset featured two types of processes - those comprising postures with durations mostly lower than 30s (short postures) and processes with postures of durations over several minutes (long postures). A process either contains one long posture sequence or 10 short posture sequences executed in a row. The mean duration of a process was around 1400 seconds or 24 minutes. As DAT1 comprised nine posture sequences, the total number of DAT3 processes was 18. Each DAT1 sequence was transformed into a corresponding DAT3 process with short postures and another with long postures. Figure 8.5 shows an example of how a posture sequence from DAT1 was converted into DAT3.



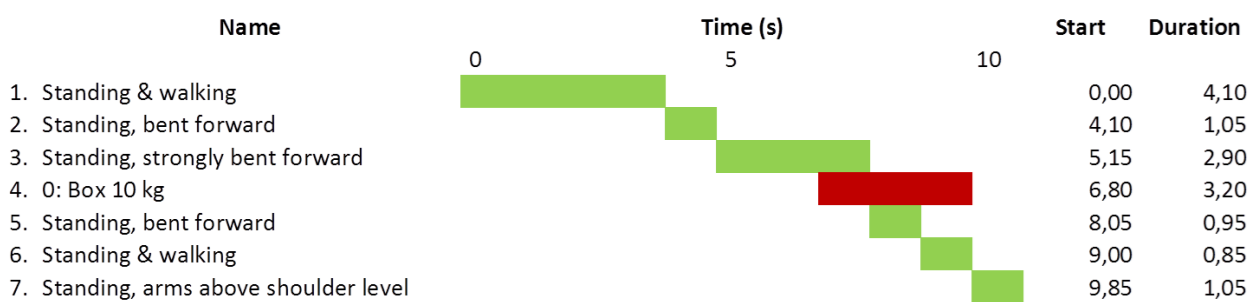
Figure 8.3: Sample frames from DAT2.



(a) Subject with markers attached



(b) Postures transferred to the DHM[70]

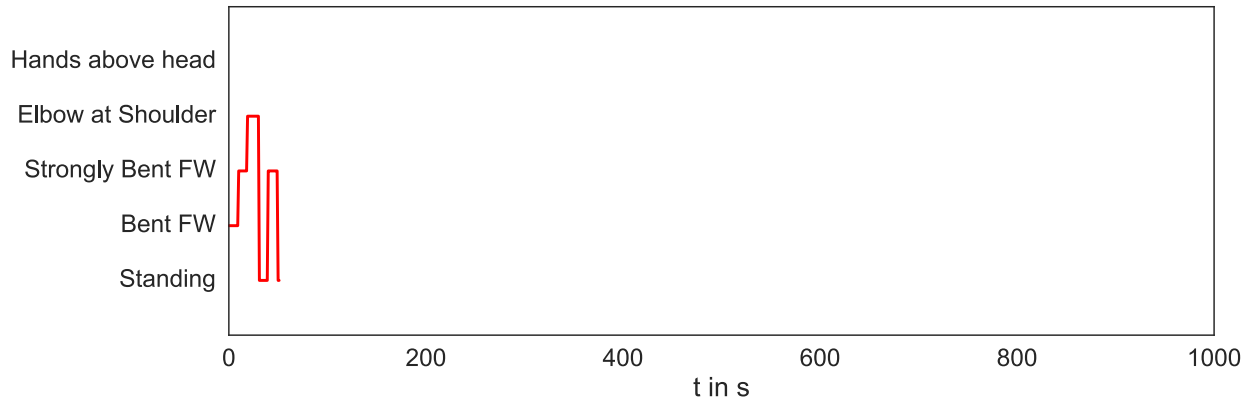


(c) Posture classification result

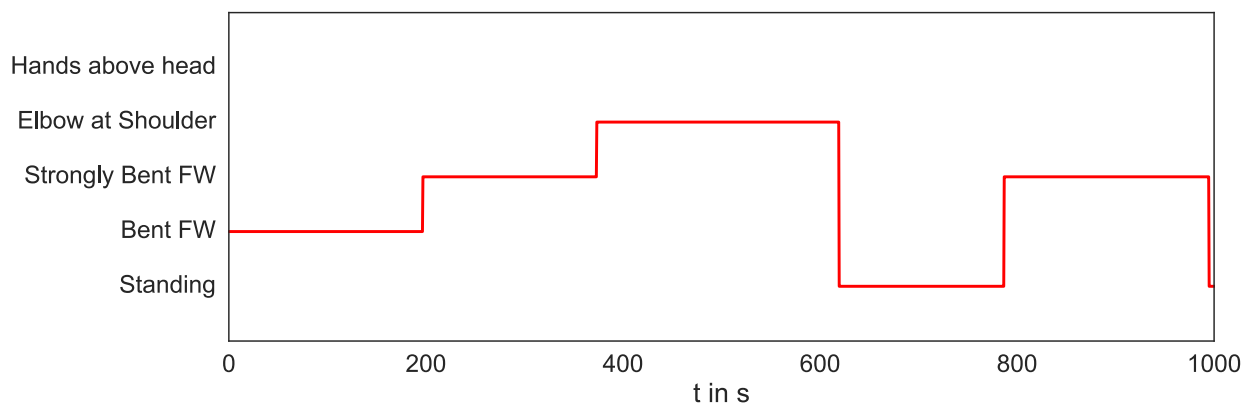
Figure 8.4: Workflow to obtain EAWS postures from subject motions.

8.2 Processing Time

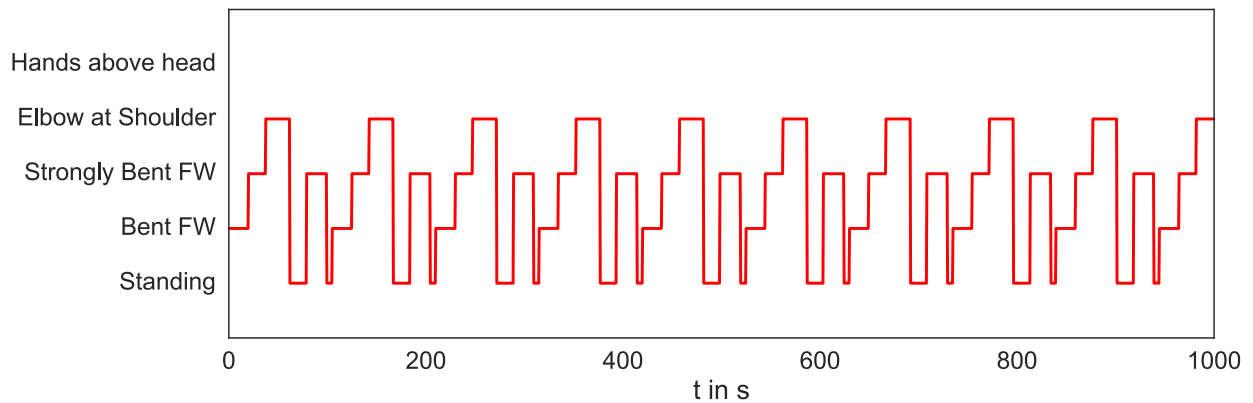
Runtime tests were conducted using the WPCApp. The test system was a laptop with an Intel Core i5 2520m CPU with two 2.5GHz cores. The system accessed 8GB RAM



(a) Sequence from DAT1



(b) Long posture process



(c) Short posture process

Figure 8.5: One exemplary posture sequence from DAT1 converted into a DAT3 sequence.

and ran on the Windows 10 operating system. The runtime measurements were conducted through a test video sequence from DAT1. As performance measures, the mean and standard deviation of the processing time for a frame were determined. A WPC processing loop includes data acquisition, pre-processing, posture assessment and posture optimisation. The two latter stages took up the majority of the processing time. Posture

assessment using the 2D localisation algorithm takes $197 \pm 69ms$, while using 3D features consumes roughly 10s per frame. Posture optimisation with a height-adjustable actuator takes $0.6 \pm 3ms$. The computation of an adjustment for the robot actuator is on average $8 \pm 8ms$ long. Hence, the posture assessment component dominates the computation time.

Result Summary

In total, the WPC processes one loop in less than 300ms. Considering that the EAWS requires a time granularity of one to four seconds [34], this performance is sufficient.

8.3 Posture Assessment

8.3.1 Standing Postures (DAT1)

First, the performance of each posture classification approach was determined. A leave-one-out cross-validation setup was used where in each round or fold, one posture sequence was employed as test set and eight sequences were utilised to train the classifier. To obtain the end result, the intermediate results from all folds were averaged. With this setup, the rather small dataset can be fully exploited. Each of the 9 sequences can be used for training as well as for testing without distorting the outcome. The comparison comprised the joint angle features computed from the 2D joint localisation, 3D joint localisation and NiTE2 joint localisation [169] algorithm. The classifiers included:

- K-Nearest Neighbours with K=5 (KNN) and Euclidean Distance
- Support Vector Machine with Linear Kernel (LIN-SVM)
- Support Vector Machine with Radial Basis Function Kernel (RBF-SVM)
- Random Forests (RF) with 10 Trees

To obtain the best SVM parameters, a grid search was applied on the training data. The Scikit-Learn library [197] provided the classifier implementations.

Posture Classification Accuracy

To quantitatively evaluate the results, the classification accuracy measure was leveraged. Let $n_{Correct,i}$ denote the number of correctly classified samples in fold, i , and $n_{Wrong,i}$ the number of wrongly classified samples, respectively. The accuracy score, acc_i , for fold, i , was computed as:

$$acc_i = \frac{n_{Correct,i}}{n_{Correct,i} + n_{Wrong,i}} \quad (8.1)$$

The overall accuracy over all N folds was computed as:

$$acc = \frac{1}{N} \sum_{i=1}^N acc_i \quad (8.2)$$

Table 8.2 lists the accuracy scores for all feature-classifier combinations tested. The results

Table 8.2: Mean accuracy and mean EAWS differences for the tested feature-classifier combinations. Best values are marked bold.

Feature	Classifier	Accuracy (acc)	Mean difference in EAWS points
2D	RBF-SVM	0.95	3.2
2D	LIN-SVM	0.90	3.8
2D	RF	0.95	3.7
2D	KNN	0.95	3.8
3D	RBF-SVM	0.44	18.8
3D	LIN-SVM	0.40	19
3D	RF	0.56	12.4
3D	KNN	0.44	13.3
NiTE2	RBF-SVM	0.29	11.3
NiTE2	LIN-SVM	0.29	9.8
NiTE2	RF	0.32	10.1
NiTE2	KNN	0.28	10.3

from the experiments suggested that the 2D features were superior for the application case in this thesis. Supportive evidence was found for this assumption by visualising each feature type with the t-distributed stochastic neighbour embedding (t-SNE) technique [198]. In contrast to the similar dimensionality reduction technique, principle component analysis (PCA), t-SNE computes a non-linear mapping, which better represents the real distances between the data points. Figure 8.6 depicts the joint angle vectors transformed into 2D space. The result plots clearly indicate that 2D features and 3D features are better separable by a classifier than NiTE2 features, where posture regions heavily overlap. Conversely to the classification results, 2D features and 3D features seem to be comparably well separable.

Interesting insight can be obtained by viewing the confusion matrices. A confusion matrix, \mathbf{C} , with the indexing \mathbf{C}_{ij} indicates how many samples of the posture class, i , were classified as class j . Figure 8.7 depicts the results for the 2D, 3D and NiTE2 features in combination with the RBF-SVM. The amounts were normalised by the total number of class i samples, meaning the value \mathbf{C}_{ij} was divided by $\sum_j \mathbf{C}_{ij}$. According to the confusion matrices, there was one dominant error category for the 2D features - "Standing" postures were often confused with "upright with elbow at shoulder level". Presumably, this error pattern corresponded to the problem that 2D features were susceptible to changes in the length of the upper limbs due to projection. The confusion matrices for the other features depicted more equally distributed error types.

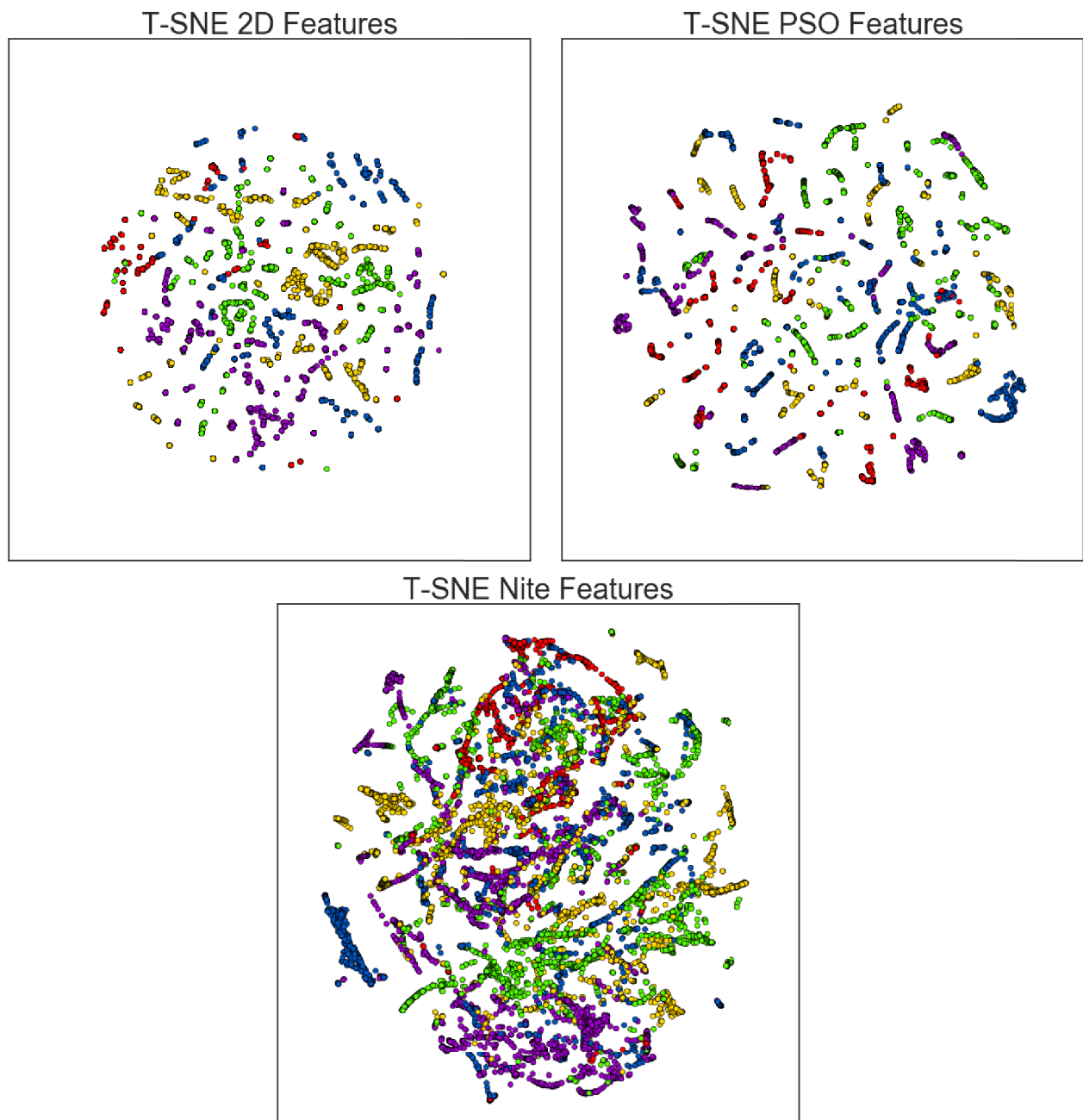


Figure 8.6: Visualisation of the feature space for different approaches using t-SNE. Different colours refer to varying class labels.

Comparison of Automatically and Manually Determined EAWS Scores

In considering the assessment of ergonomics, the classification accuracy itself had limited significance. If two postures which shared similar EAWS scores are confused, the assessment will not be severely affected. The same cross-validation setup as described earlier was used, but the EAWS was directly computed on the classified posture sequence. The fourth column of Tab. 8.2 outlines the mean of absolute differences of EAWS scores between the automated posture assessment algorithm and the manually determined scores. For a more detailed view, the EAWS differences for each video sequence were visualised

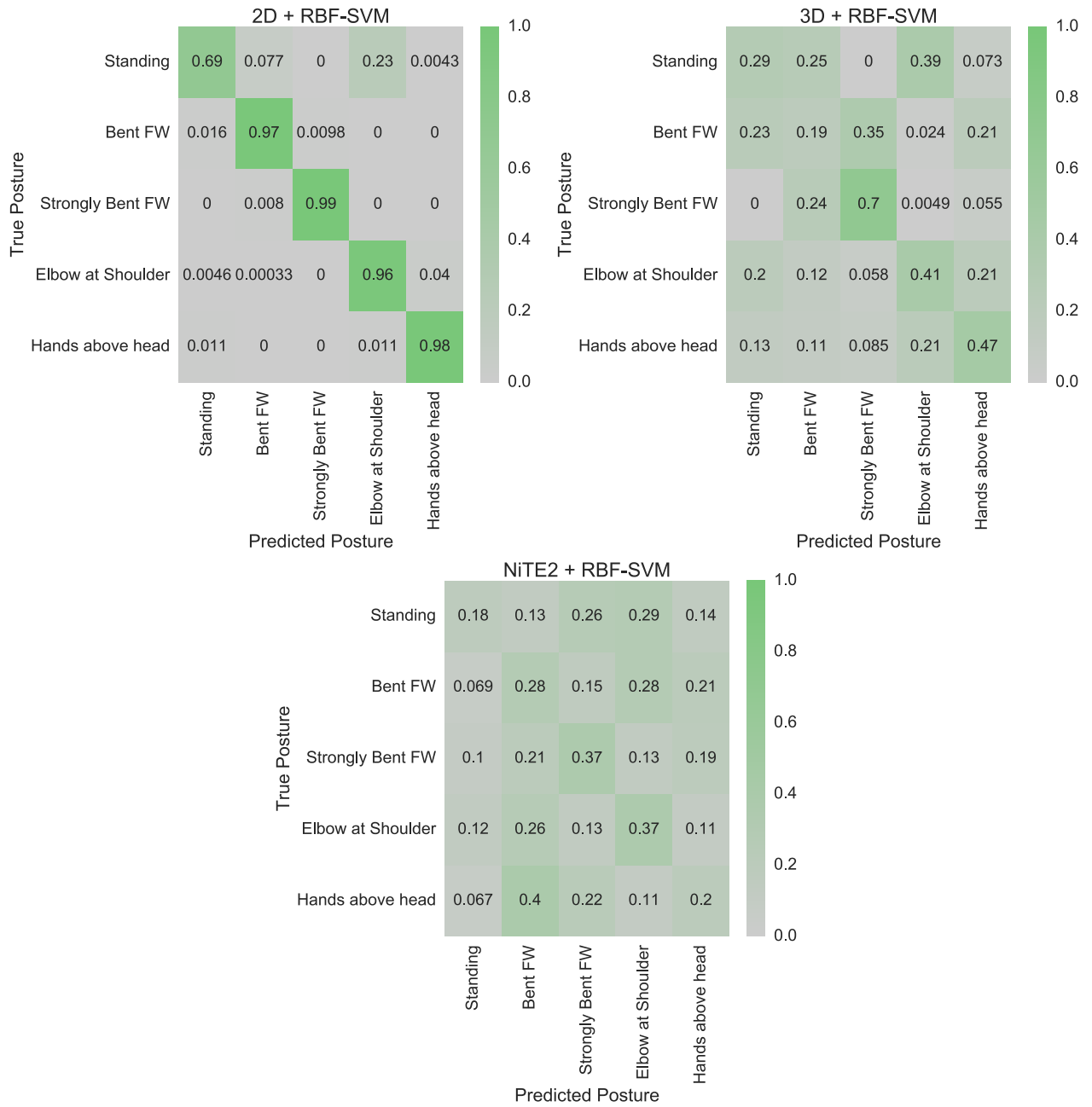


Figure 8.7: Confusion matrices achieved with joint angle features from different localisers.

for an exemplary feature-classifier combination in Fig. 8.8. In order to better understand the units of the EAWS score difference, consider the risk classification scheme of the EAWS (see Sec. A of the Appendix). Tasks yielding 0 to 25 points are classified as low risk. If the EAWS score value is between 25 and 50 points, the task is considered as possibly risky. Finally, EAWS values above 50 points indicate a task from the high risk category. Thus, a low one-digit EAWS difference value unlikely leads to a misclassification of a task whereas higher mean differences increase the likelihood of classification errors.

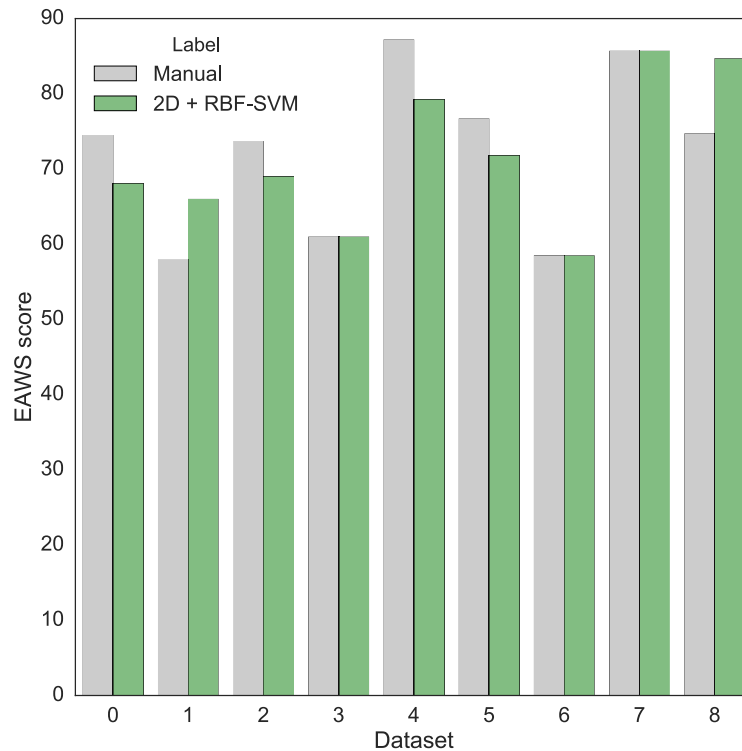


Figure 8.8: EAWS score comparison for the 2D + RBF-SVM combination.

Result Summary

According to the results, 2D features yielded the best classification accuracies as well as the most precise EAWS computations by high margin (see Tab. 8.2 first row). Figure 8.7 supports this statement. In the "2D + RBF-SVM" case, the highest values are concentrated on the main diagonal of the confusion matrix. In contrast, the other feature types result in values distributed all over the matrix. 3D features produced better classification accuracies than the ones computed from the NiTE2 library (see Tab. 8.2 third column, rows 5-8 versus rows 9-12), however the classification errors had a more severe impact on the EAWS accuracy versus NiTE2 features. Finally, it could be seen that the best classifiers for these features are RBF-SVM and RF. Whereas both offer comparable results on the 2D features, RF slightly outperforms RBF-SVM on 3D and NiTE2 features. In terms of run-time, all the algorithms using 2D features or features from the NiTE2 library are able to output classifications in roughly 300ms. When using 3D features, most of the computation time is dedicated to the feature extraction step, which takes up to 10s per frame on a state-of-the-art GPU whereas classification is performed in less than 200ms. If feature extraction is not included into training time, all proposed algorithms can be trained on the whole dataset of around 6000 data points in less than 3 minutes.

8.3.2 Comparison to Alaska/EAWS Assessment (DAT2)

Experiments on DAT1 indicated that 2D body angle features in combination with an SVM classifier (2D + RBF-SVM) performed best. Its classification accuracy was investigated more in detail via DAT2. Testing other feature-classifier combinations was skipped due to the clear results on DAT1. As a result of the higher number of labels, the classification problem here was expected to be more challenging than in DAT1.

Comparison of 2D + RBF-SVM and Manual Posture Classification

As mentioned in Sec. 8.1, each video in the DAT2 dataset contained only one posture class. Thus, there was one manually determined posture label per video sequence. To obtain the 2D + RBF-SVM posture labels, the posture classifier was applied on each frame and the most frequently occurring label was computed. In comparison to the manual labelling, the developed algorithm achieved an accuracy of 75%. The resulting confusion matrix is depicted in Fig. 8.9. As can be observed, most errors occurred with sitting and kneeling postures, especially when the upper limb posture was "elbow at shoulder level". A possible interpretation is that the criteria for determining this posture are less clearly stated in the EAWS than for other postures. In particular, no angle intervals are provided for the upper limb posture classification. Therefore, the way subjects perform this posture is more variant than in other cases, making a precise classification more challenging.

Comparison of 2D + RBF-SVM and Alaska/EAWS Posture Classification

The classification performance of the posture assessment algorithm was compared with posture classifications from the commercial DHM software, Alaska/EAWS [70]. For both classifiers, the most frequently occurring posture of one sequence was set as the video label. The resulting confusion matrix is found in Figure 8.10. When using Alaska/EAWS as a ground-truth label, the classification performance of the algorithm decreased to 47%. At first glance, it stands out that Alaska/EAWS rarely classified upper limb postures as "elbow at shoulder level". A discussion with the software support team revealed that the software selects this upper limb posture when only one arm is at shoulder level. As the subjects performed symmetric upper limb postures, this label was technically not included in the dataset. Furthermore, back postures classified as "bent forward" by Alaska/EAWS are often misclassified by the algorithm developed in this thesis. However, these postures are also the ones where manually created labels and the labels generated by Alaska/EAWS were in least alignment. A possible explanation is that when bending the back forward, the visibility of certain markers is impaired, leading to misclassifications of Alaska/EAWS. In fact, both ground-truth labels only coincided in 19 out of 32 cases (58%). Figure 8.11 presents the confusion matrix between manually determined labels and those labels determined by Alaska/EAWS.

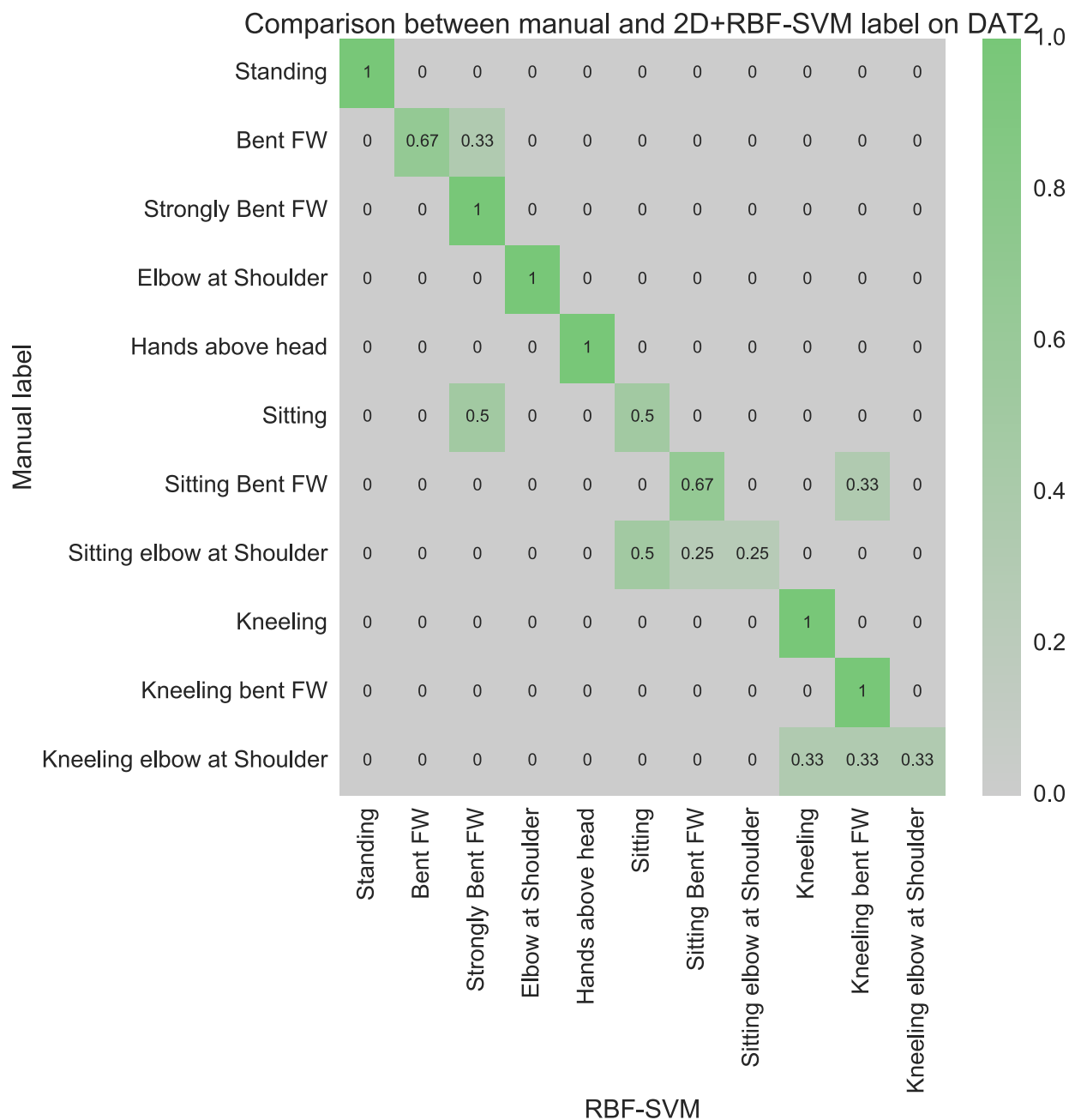


Figure 8.9: Difference between 2D+RBF-SVM labelling and manual labelling.

Result Summary

In comparison to DAT1, the posture assessment algorithm performed worse on DAT2. In particular, when employing the Alaska/EAWS labelling as ground truth, the classification accuracy was dramatically diminished. This can be partially explained by the lack of marker visibilities in the "bent forward" postures, leading to erroneous classifications by Alaska/EAWS. However, it was also observed that manual and Alaska/EAWS labellings strongly diverged. In conclusion, the feature set created for posture assessment was strongly tailored to the own understanding of the postures. Adapting the algorithm

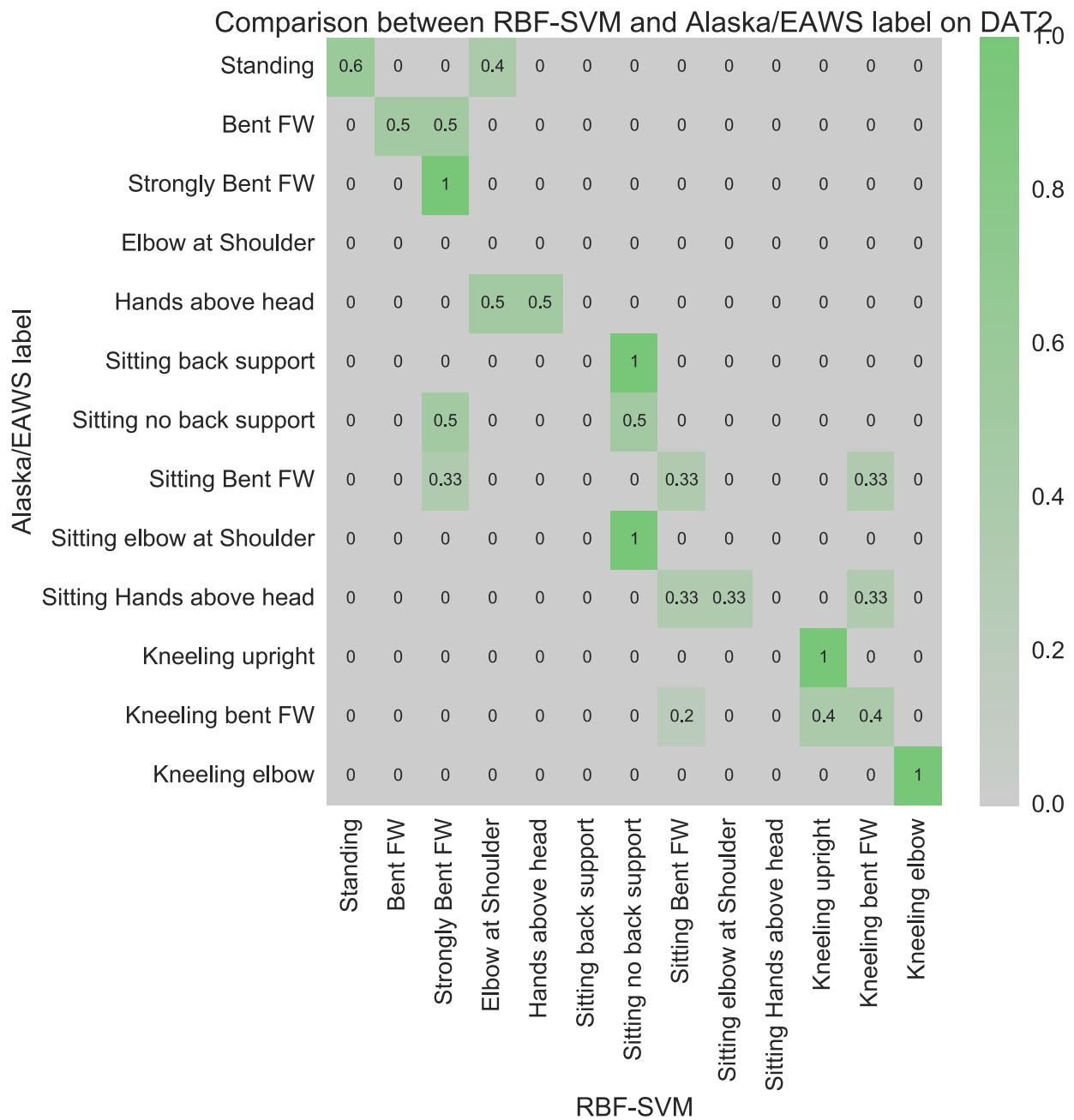


Figure 8.10: Difference between 2D+RBF-SVM labelling and Alaska/EAWS labelling.

to the Alaska/EAWS ground truth would require a fundamental re-design of the feature extraction pipeline.

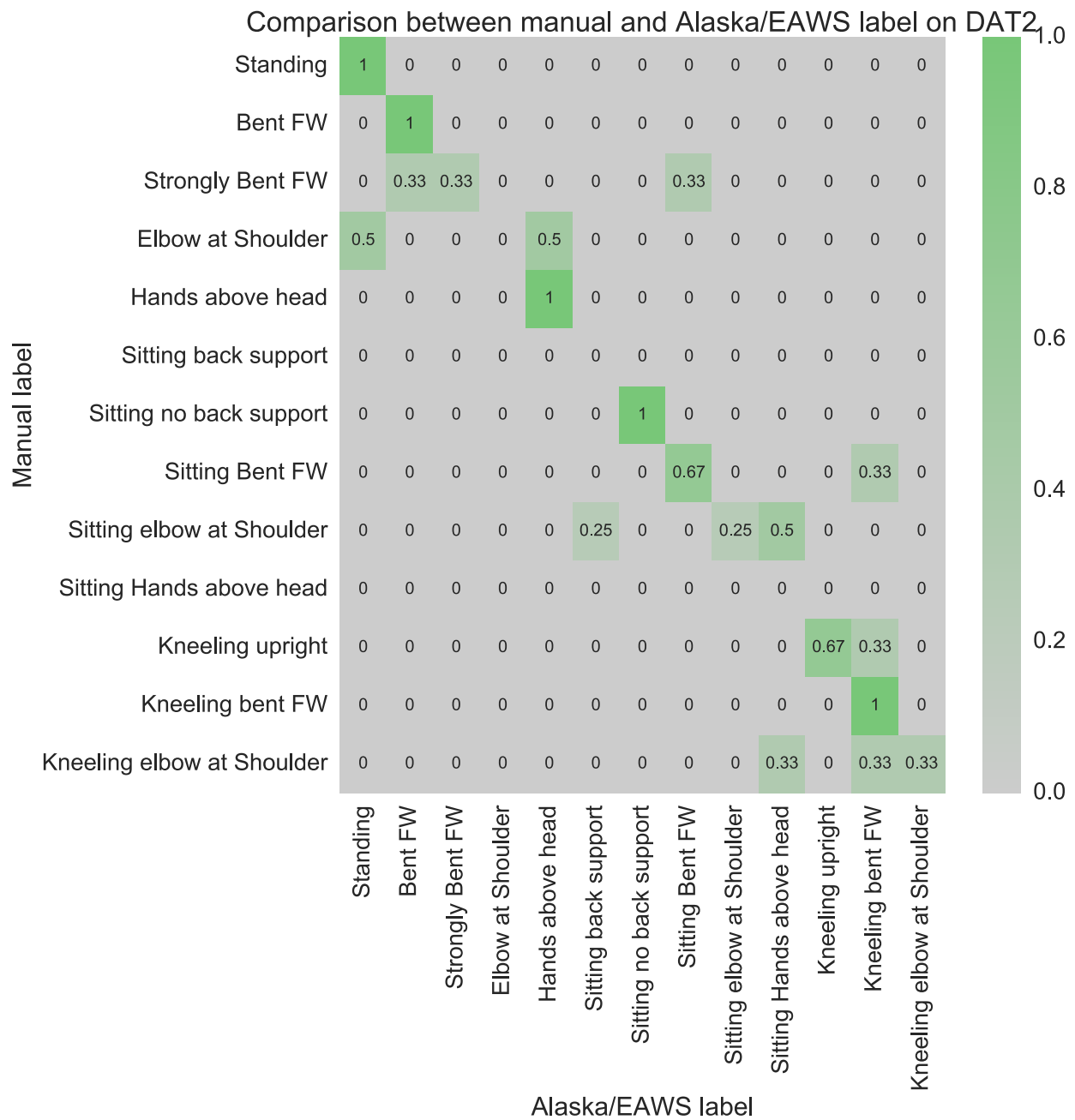


Figure 8.11: Difference between manual labelling and Alaska/EAWS labelling.

8.4 Posture Optimisation

8.4.1 Comparison of Height-Adjustable Actuator and Robot Actuator (DAT1)

With DAT1, the performance of the posture optimisation component was evaluated. First, the workpiece pose search routine was analysed. Posture optimisations were run on all frames of each video containing a human subject. Figure 8.12 shows selected frames containing the predicted posture. The red human model visualises the original posture. The

green human model denotes the posture when employing a height-adjustable actuator. The cyan human model presents the resulting posture with a robot actuator.

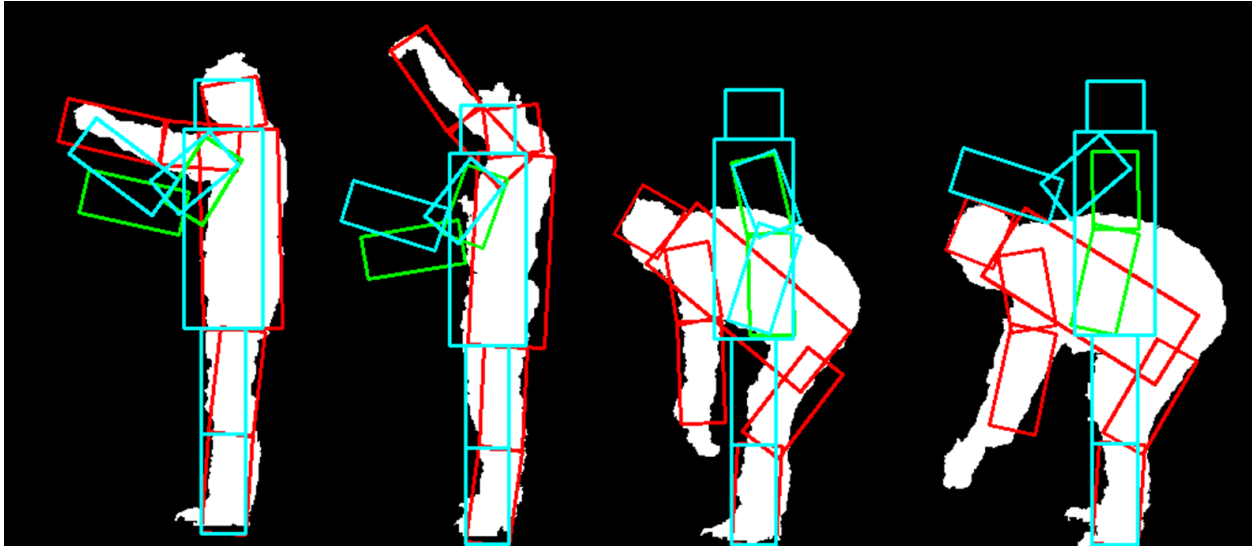


Figure 8.12: Qualitative posture optimisation results.

Posture Distribution Before and After Optimisation

To obtain quantitative results, the distribution of posture labels before and after optimisation was analysed. Specifically, the number of cases of a particular label that had been transformed into another, was determined. Figure 8.13 describes the statistics.

In contrast to the stated expectations, the results suggested that the height-adjustable actuator led to more robust ergonomic results than the robot actuator. In particular, the robot actuator did not manage to improve postures where the arm was raised. One interpretation is that the posture optimisation problem formulated for the robot actuator is more complex, resulting in problems with local minimum solutions. In the robot actuator optimisation problem (see Eq. 6.27), the actuator configuration, \mathbf{p} , had to be optimised, which can be determined in a simpler way for the height-adjustable actuator. However, additional analyses in this section (see Fig. 8.16) suggest that the better postures achieved by the height-adjustable actuator can be caused by an anomaly case of the algorithm. The optimisation of the height-adjustable actuator was trapped into a local minimum right at the point of initialisation, which was defined as the ergonomically ideal posture. To sum up, non-linearities from the kinematics of human model and actuator lead to a complex optimisation problem with a high number of local minima in both cases. This problem can be mitigated by a routine which outputs an initialisation depending on the posture before optimisation. Machine learning approaches could determine the mapping between posture before optimisation and appropriate initial parameters from data.

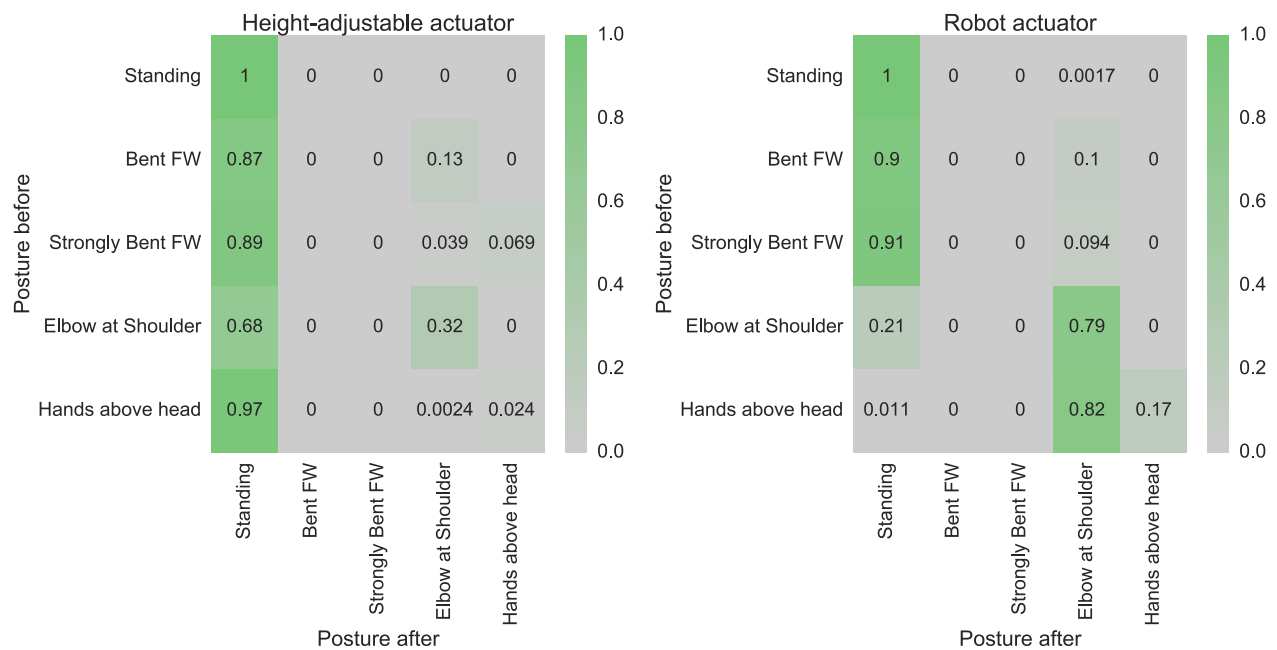


Figure 8.13: Distribution of postures before and after optimisation for height-adjustable and robot actuator.

Difference in Working Direction

The motivation for using a robot actuator instead of a height-adjustable actuator was that it would enable tilting of the workpiece. The assumption behind was that additional DoF allowed better maintaining of the original working direction. Thus, the robot would better enable the worker to perform the task at hand. As an indicator of whether the worker could perform this task after optimisation, the difference between working direction and workpiece orientation before and after adjustment was measured. In theory, the lower this angle difference was, the better the worker could perform the task as the workplace layout was more similar to the original.

The distribution of working direction angles before posture optimisation (original working direction) is shown in Fig. 8.14. The angle values roughly ranged from -250° to 0° . Figure 8.15 outlines the mean absolute working direction difference as a function of the original direction. Both graphs have a U-like shape indicating that the actuator algorithms work best in the angle interval between roughly -150° and -70° .

Contrary to the initial intuition, the graphs imply that the height-adjustable actuator better maintains the original working direction in most cases. Only if the original working direction is higher than -90° , the robot actuator is a better choice. Angle values higher than -90° occur when the subject has to access the workpiece from below. The only way for a height-adjustable actuator to enable processing the workpiece from below is by raising the working height above the torso. However, this leads to postures where arms are required to be located at shoulder level or even above head level, which are ruled out by the algorithm due to unacceptable ergonomic costs. The robot actuator is able to slightly

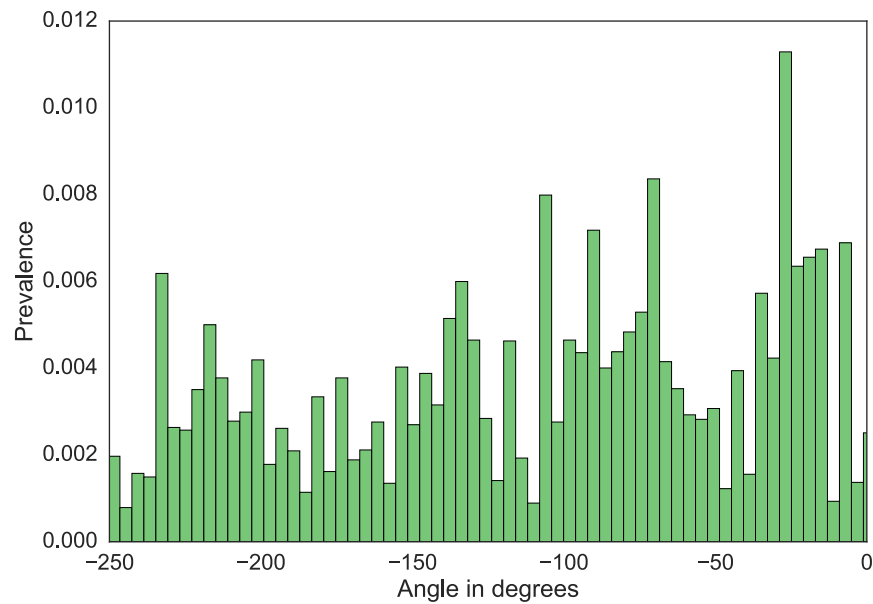


Figure 8.14: Distribution of working direction angles in DAT1 before posture optimisation.

better maintain the working direction difference through its ability to tilt the workpiece. Nevertheless, the working direction maintenance still poses significant ergonomic costs (see Fig. 8.13). Presumably, the superior performance of the height-adjustable actuator for angle values below -90° is because of the more complex optimisation problems needing to be solved for the robot actuator. Especially omitting the contact point objective helps to reduce the difficulty of the optimisation.

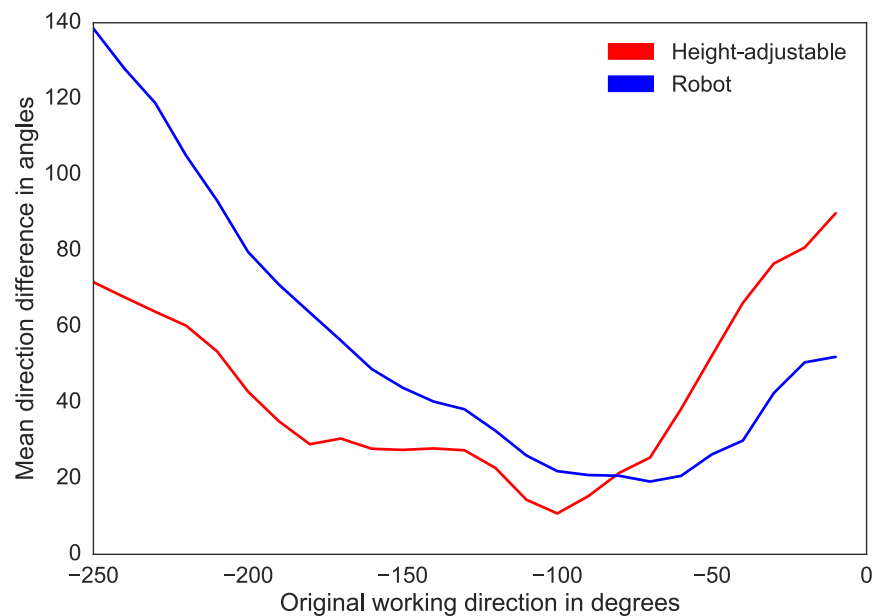


Figure 8.15: Mean absolute difference of working directions as a function of the original working direction value.

Figure 8.16 provides a more in-depth insight into the error structure. It expresses the distribution of working direction angles before and after optimisation. Each working angle combination is represented by one data point. Working direction angles before optimisation are represented by the x-value. The point's y-value denotes the corresponding angle after optimisation. An algorithm perfectly maintains the working direction if all of its data points are located along the visualised diagonal line. There is an anomaly case where both algorithms, independently from the original direction, output a posture with a working direction at around -100° (see region between dashed horizontal lines). These working directions correspond to the initialisation parameters of the optimisation routine. It is conjectured that the present implementation of the SQP algorithm from the NLOpt library gets trapped into a local minimum near the initial parameters. Besides, the figure suggests that the height-adjustable actuator hardly outputs configurations where the optimised working direction is above -90° .

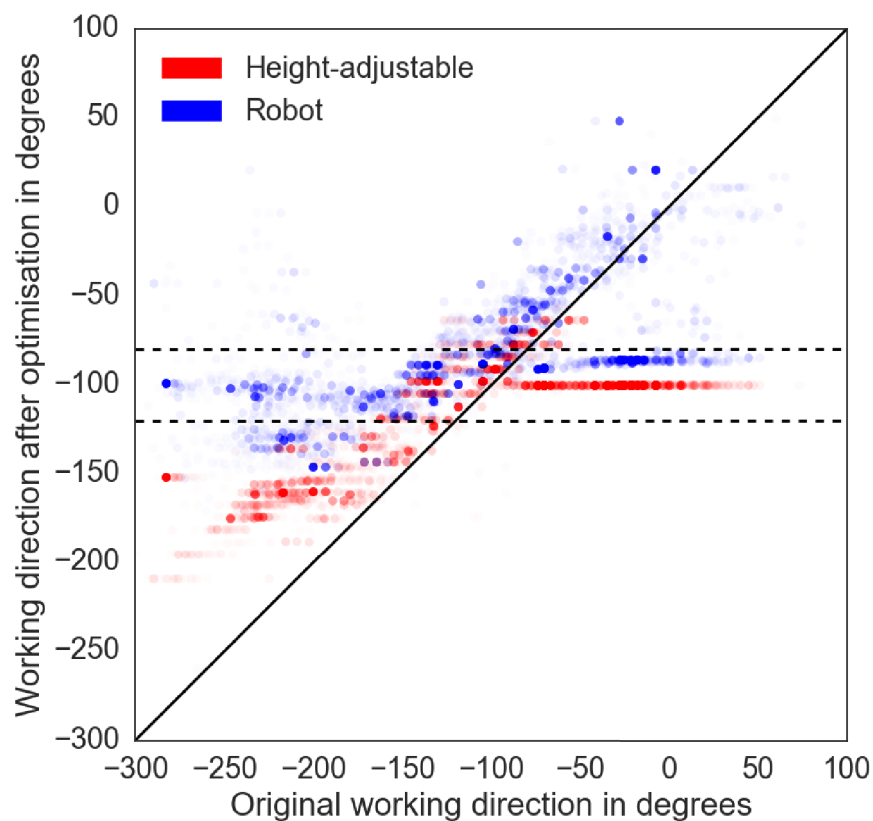


Figure 8.16: Distribution of working direction values before and after optimisation. The lower the transparency of the data points the higher the density.

Result Summary

The posture optimisation algorithm managed to find more ergonomic postures especially for cases where the worker would need to bend forward (see Fig. 8.15). Despite having fewer DoF, the algorithm for the height-adjustable actuator outperformed that for the robot actuator in terms of working direction maintenance. For cases where the workpiece

is located too high, the results are more controversial. The robot actuator with its tilting ability had the potential to better support task accomplishment (see Fig. 8.15). However, this benefit comes with significant ergonomic costs (see Fig. 8.13). In general, both algorithms struggle with the problem that there is an undesired local minimum near the initial parameters, which becomes the outputted solution (see Fig. 8.16).

8.4.2 Comparison of TET and DTET Control Strategies (DAT3)

The performance of the developed control strategies was evaluated on dataset DAT3. Simulation experiments were employed to determine the performance of each control strategy and its parameters. Simulations have the advantage that they enable extensive testing with a variety of parameter sets with reasonable time effort. Furthermore, the ergonomic score and the overall ADAP can be recorded without additional effort. On the other hand, a simulation does not replace tests with subjects after development of the workplace.

Simulation Environment

The WPC components (control strategy, posture assessment and workpiece pose search) were integrated into a testbench (see Fig. 8.17 for overview). Human behaviour was approximated by a component that attempts to predict the worker's posture for the current task depending on given actuator parameters. The newly predicted posture triggers a loop, resulting in an actuator parameter update. Besides this, the EAWS score and the ADAP are recorded. The following paragraphs briefly describe each element of the testbench:

Consider the sequence of postures adopted when the actuator parameter remains at the initial height. These postures are referred to as original postures, representing the worker's behaviour when there was no adjustment at all. From an original posture in a frame, the contact point, $\mathbf{x}_{Contact}$ (see Fig. 6.9 for visualisation), was established. Having computed the current contact point, $\mathbf{x}_{Contact}(t)$, for frame, t , the component passes the original posture parameters, $\mathbf{q}_{orig}(t)$ and $\mathbf{x}_{Contact}(t)$, to the posture prediction component.

The posture prediction component simulates human behaviour by predicting the posture angles, $\mathbf{q}'(t)$, from the current actuator parameter, $\mathbf{p}'(t)$, and the original posture parameters, $\mathbf{q}_{orig}(t)$ and $\mathbf{x}_{Contact}(t)$. The approach used is similar to that described in Sec. 6.3. In the case of the height-adjustable workplace, the predicted body angles are the outcome

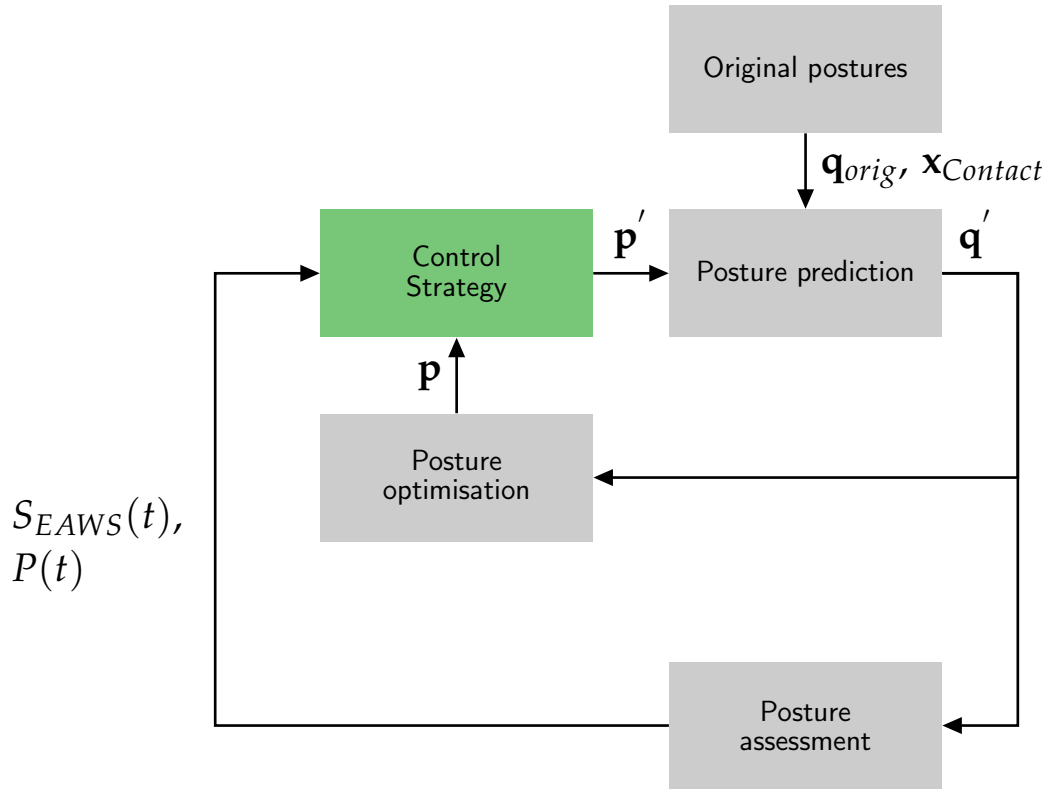


Figure 8.17: Interaction between the testbench components.

of solving the optimisation problem akin to Eq. 6.19:

$$\begin{aligned}
 \mathbf{q}' &= \arg \min_{\mathbf{q}} \\
 &\quad \|\mathbf{q} - \mathbf{q}_{ideal}\|^2 \\
 &\quad + (\mathbf{x}_{Foot,x} - \mathbf{x}_{Centroid,x})^2 \\
 &\quad + \left(\sum_j \mathbf{q} - \sum_j \mathbf{q}_{Before,j} \right)^2
 \end{aligned} \tag{8.3}$$

subject to

$$\begin{aligned}
 \mathbf{q}_{min} &\leq \mathbf{q} \leq \mathbf{q}_{max} \\
 h_{min} &\leq \mathbf{x}_{Platform}(\mathbf{p}) + \mathbf{x}_{Hand}(\mathbf{q}) - \mathbf{x}_{Hand}(\mathbf{q}_{Before}) \leq h_{max} \\
 \mathbf{x}_{Platform}(\mathbf{p}) + \mathbf{x}_{Contact} - \epsilon &\leq \mathbf{x}_{Hand}(\mathbf{q}) \leq \mathbf{x}_{Platform}(\mathbf{p}) + \mathbf{x}_{Contact} + \epsilon
 \end{aligned}$$

The posture prediction problem formulation, Eq. 8.3, adds an additional constraint to the original problem (marked in red) by forcing the position of the hand, $\mathbf{x}_{Hand}(\mathbf{q})$, to remain in a neighbourhood of the computed contact point, $\mathbf{x}_{contact}$. This ensures that the worker will opt for a posture that enables processing the same contact point.

The posture assessment component is identical to that described in Chapter 5. The posture classifier uses the joint angle features, $\mathbf{q}'(t)$, from the posture prediction component. The posture assessment component outputs the temporal EAWS $S_{EAWS}(t)$ and the currently classified sequence of postures, $P(t)$.

The posture optimisation component was equal to that from Sec. 6.5. The predicted posture, \mathbf{q}' , was set as the input posture and the input actuator parameters were obtained from the last control strategy step, $\mathbf{p}'(t-1)$. The predicted posture of the posture optimisation component was discarded as it did not take the contact point of the next frame $t+1$ into consideration. Rather, the predicted posture, $\mathbf{q}'(t)$, was used for the next loop.

Finally, from the temporal EAWS score $S_{EAWS}(t)$ and the posture sequence, $P(t)$, the control strategy pursued whether to utilise the proposed actuator adjustment from the posture optimisation component or not. It outputs the signal, d_{adj} .

Test Setup

The control strategies were evaluated in combination with a height-adjustable actuator. Based on the fact that the control strategy did not depend on actual actuator configurations, but only on the ergonomic scores, experiments were strictly conducted with the height-adjustable actuator without losing generality. Two configurations of the actuator were simulated. They differed in terms of adjustable height ranges. A small range actuator was able to set the platform height between $50\text{cm} - 150\text{cm}$, while a hypothetical large range actuator was set to implement platform heights between $0\text{cm} - 200\text{cm}$. The actuators do not correspond to real world devices. Rather, the purpose of these experiments was to theoretically investigate influence of the range on the quality of the posture optimisation.

The TET and DTET control strategies were appraised on each of the DAT3 processes, recording the EAWS score as well as the ADAP in cm. The experimental setup assumed a 50 percentile human according to DIN 33402-2 [168]. The actuator height at the beginning of the simulation was set to 80cm .

The experiments were carried out adhering to different control strategy parameters. For the TET strategy, T was increased from 20 to 90 points in 10 point steps. For the DTET strategy, t_{Total} varied from $20\% - 100\%$ of the total process time. In each of the DTET experiments, the parameters $T_{Start} = 20$ and $T_{End} = 100$ were assumed.

Overall Control Strategy Performance

For each combination of dataset, posture length (short and long) and standard actuator configuration (none, mean height, median height), the final ADAP and EAWS score were determined. In total, 180 data points (10 posture sequences \times 18 parameter and standard actuator configuration combinations) were obtained. Out of these data points, the pareto front was computed by removing the dominated points. Dominated data points are cases where there was another data point with a lower EAWS as well as ADAP. The number of non-dominated data points belonging to the TET and DTET strategy were then determined. In all cases except one, the TET strategy had a higher contribution to the pareto front than the DTET strategy. Table 8.3 compared the numbers of non-dominated points between the TET and different variants of DTET strategies.

Table 8.3: Number of non-dominated TET and DTET data points for various posture length and standard height strategies.

Posture length	standard actuator configuration	TET	DTET
Short	none	37	32
Short	mean height	36	34
Short	median height	34	38
Long	none	37	32
Long	mean height	54	36
Long	median height	48	32

This result does not necessarily mean that the DTET strategy is inferior. When looking at the ADAP-EAWS graph portraying the non-dominated points in all experiments (see Fig. 8.18), the regions with low ADAP ($< 100\text{cm}$) or high ADAP ($> 1500\text{cm}$) appear to be dominated by the TET strategy. However, in the central ADAP regions, where an acceptable tradeoff between EAWS and ADAP ($250\text{cm} - 1000\text{cm}$) lies, there are mostly DTET data points. In other words, if the application requires a high amount of intervention to keep the ergonomics score as low as possible or if the intervention should be kept at a minimum level, then the TET strategy is recommended. The DTET strategy, meanwhile, is better for achieving compromised settings in between.

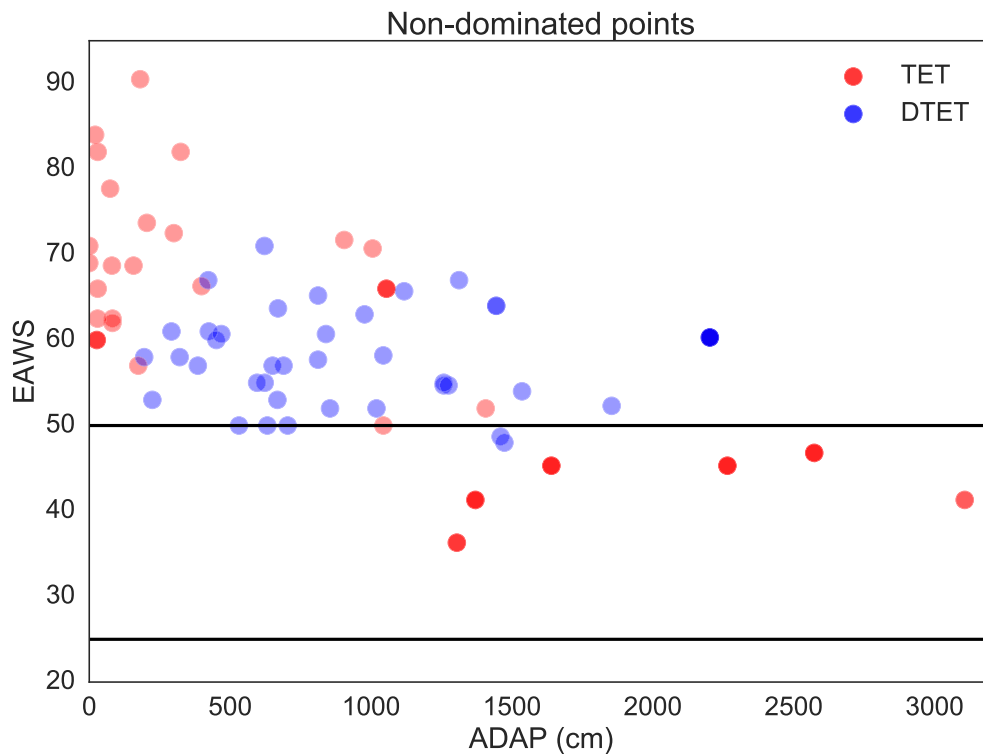


Figure 8.18: Distribution of non-dominated points.

Influence of Control Strategy Parameters

The TET strategy clearly poses a trade-off between final EAWS score and ADAP based on the temporal EAWS threshold parameter, T (see Fig. E.1a in the Appendix). The more often the actuator was adjusted, the lower and thus better the EAWS score was. This observation agrees with the intention of the control strategy described in Sec. 6.8.1. In tasks where appropriate working heights significantly vary over time, such as in the DAT3 dataset, adjusting the actuator for each new posture enables the worker to more often adopt ergonomically favourable postures. Hence, frequent adjustments lead to a low final EAWS score. However, there were exceptions - in a number of cases, a worse EAWS score was achieved despite a lower T . The problem lies in the calculation scheme of the EAWS (see Appendix A for a detailed explanation). Whereas the relationship between ADAP and EAWS can be clearly observed for the TET strategy, this tendency is not as present for the DTET strategy (see Fig. E.1b of the Appendix). The relationship is more readily apparent when applied to the short posture processes versus the long posture sequences. It should be emphasised that the posture sequence after t_{Total} heavily influences the outcome of the experiment. When the postures with high load are present after t_{End} , they cannot be treated, leading to a significant EAWS score increase.

Influence of Actuator Range

In general, the actuator range has an impact on the achievable EAWS when the control strategy encourages frequent adjustments (see Fig. E.2 of the Appendix). With the TET strategy, especially the lower temporal EAWS thresholds profit from a higher range. The DTET strategy, contrastingly, does not profit as well from a higher range as the number of adjustments is reduced. The larger the range of possible actuator configurations, the higher the chance of achieving superior ergonomic results. Yet, especially when commonly used, the ADAP increases.

Comparison of Different Posture Durations

In general, both control strategies accomplished similar results independent from posture duration. The only difference was that the ADAP was significantly higher for short postures when the processes endured the same amount of time. Naturally, this came from the need for more frequent adjustment of the workplace. A representative result is depicted in Fig. E.3 of the Appendix.

Result Summary

To sum up, a TET control strategy with low T is advised when the goal is to minimise the EAWS score. Alternatively, the TET strategy can be applied to minimising ADAP by instituting a high T value. However, the relationship between T and the EAWS score are not perfectly monotonic because of the computation scheme of the EAWS, which penalises a high share of medium-load postures more than a low share of high-load postures. When searching for compromises between EAWS and ADAP, the DTET strategy is

a better choice than the TET strategy. In addition to the control strategy, one of the most influential factors on the EAWS is the actuator range.

8.5 Conclusions

This chapter has experimentally evaluated the components of the WPC. Three datasets to assess each relevant element have been presented. The runtime analysis has demonstrated that even on a standard laptop, an evaluation loop can be realised in less than 300ms, which is sufficient for the stated application case.

The posture assessment experiments compared the various joint localisation approaches and their influence on the assessment results. As ground truth, manual assessments have been compiled. The results uncovered that features from 2D joint localisation in combination with a non-linear classifier were most suited for this application. 2D features even surpass the standard localisation implementation for the Kinect [169]. In comparison to the commercial software Alaska/EAWS, which uses the posture data from a Vi-con system, the developed algorithm significantly disagrees in the posture classification. However, the Alaska/EAWS results also differ from the manual posture labellings. In general, the quality of the ground truth data turns out to be an issue. Due to the errors produced by the motion capture system, it cannot be guaranteed that the Alaska/EAWS labellings are completely correct. The manual labelling can be distorted by the fact that the labeller was not an expert. The only way to produce high quality ground truth data is having one or multiple trained EAWS practitioners, which were not available during this project. The scientific risks stated in Sec. 5.2 have turned out to be mostly resolvable. Only the question whether joint angles suffice as posture classification feature cannot be satisfactorily answered, as seen by the contradictory results between manual assessment and Alaska/EAWS assessments.

Experiments on the posture optimisation component demonstrated that the developed algorithm was able to transform most unfavourable postures into those that would be more favourable, especially when the workpiece was to be accessed from above. Contrary to stated expectations, the robot actuator algorithm fared worse with respect to ergonomics as well as the maintenance of the working direction. However, additional degrees of freedom can simplify accomplishing the original task when the object has to be accessed from below. As stated in Sec. 6.2.1 and 6.2.2, local minima turned out to be an unsolved problem, which impede the quality of the adjustments.

Finally, the dynamic behaviour of the WPC was analysed with different control strategies. As it is tedious to evaluate multiple setups using human subjects, a testbench was proposed to conduct evaluations through simulation. These simulated results revealed that the DTET control strategy is the method of choice when searching for a compromise between ergonomics and intervention frequency. The TET strategy is appropriate when ergonomics has an essential role. Furthermore, the influence of the strategy parameters,

average posture durations and actuator ranges on the results was discussed. In conclusion, the scientific risk described in Sec. 6.2.3 can be partly discounted, as the tested case only considers a cyclic task.

9 Conclusions and Outlook

This chapter reviews the work of this thesis and outlines the achievements as well as shortcomings. First, the golden thread of this thesis is summarised. The sections go over what has been accomplished and where the limits were. Thereafter, this chapter fast-forwards in an attempt to predict the impact this work could have on other areas of research. In particular, open questions that can potentially be of material for future research projects, are covered.

9.1 Summary

The starting point of this thesis was the prevalent problem of WMSDs at industrial workplaces, one of the most harmful work-related disorders. WMSDs heavily affect employees as well as employers. They are the main cause for worker absenteeism and lead to costs at the magnitude of hundreds of billions of Euros. Research has shown that one major reason for such prevalence of WMSDs is poor workplace layout design, forcing the worker to adopt awkward postures to accomplish the task at hand.

As a consequence of its significance, there has been extensive work on reducing the effect of WMSDs. A literature research on existing solutions has revealed a fundamental problem - either the solutions proposed are affordable but ineffective or they are effective albeit costly. Therefore, the access to effective WMSD prevention is restricted. This dilemma is called the "effectiveness-flexibility trade-off".

To extend the state of the art of WMSD prevention, the goal of this thesis was to bridge this trade-off. The hypothesis of this work was that it is possible to develop an intelligent system that automatically performs the costly work of specialised ergonomists in a few seconds, resulting in an effective as well as flexible solution.

The WPC, a novel type of equipment to eliminate, or at least reduce, postural load at work, has been presented. The core idea is to automate the tasks of posture assessment and workplace layout planning and optimise their runtime such that the system is able to analyse the situation and produce a solution within a few seconds. A workflow has been proposed for the WPC system that ensures a safe and efficient interaction with the worker.

Afterwards, the algorithmic manifestations of each component have been described. To realise automated posture assessment, two joint localisation algorithms based on the generative approach were developed. A feasibility analysis of potential image sensors led to the Microsoft Kinect v1 as the input device of choice. The localisation algorithms use Kinect depth images as input and compute the posture parameters. To speed up the parameter estimation, algorithmic modifications as well as parallel computing, were applied. Classifiers were trained to predict the worker's posture from the extracted parameters. Based on the sequence of occurring postures, a simplified version of the EAWS was computed.

To attain the posture optimisation component, an approach was developed that transforms the layout design task into a non-linear optimisation problem solvable by the SQP algorithm. Alternative problem formulations were discussed and tailored to operate on a height-adjustable actuator as well as a robot actuator. Furthermore, the statement was made that frequently adjusting the workplace, although ergonomically sound, results in unacceptably frequent interruptions of the workflow. Therefore, a rule-based control strategy has been proposed to reduce the postural load and simultaneously keep the adjustment activity at a minimum. After elaborating upon the concepts behind each component, the own prototypical implementation of the WPC has been presented. Special focus has been put on the design of a software architecture to promote clarity, re-usability and extensibility. The system was designed to be reconfigurable for future application cases.

Finally, the effectiveness of the algorithms under the hood of the WPC was evaluated on video datasets as well as in simulations. Experiments were performed and discussed on the aspects system runtime, posture assessment accuracy and posture optimisation effectiveness to discern whether the original goals of this thesis were reached.

9.2 Achievements of this Thesis

A novel way to prevent awkward postures at work has been presented. In contrast to approaches presented to date, the developed solution is able to solve the effectiveness-flexibility trade-off using intelligent sensors, motion analysis algorithms and actuators. A workflow of the system has been proposed that would make certain there is safety in communication between the worker and the system. The system is capable of automatically performing ergonomic assessments and provides immediate feedback. Moreover, an image-based method to perform ergonomic posture assessments has been developed. In contrast to available solutions, it does not require bulky motion capture equipment or inertial sensor systems to be worn. Furthermore, it is able to operate with a single depth camera. Also, a theoretical framework has been devised to compute a workplace adjustment that would permit the worker to carry out tasks in a more health-preserving manner. This framework was specified for a height-adjustable platform and a robot manipulator.

The evaluations showed that the joint localisation algorithm outperformed the commercially available software in the case of ergonomic assessment using the EAWS. Experiments on posture optimisation demonstrated that the algorithms find a layout that makes it possible to accomplish the original task in an ergonomically sound manner. In order to arrive at what was achieved, the research questions posed in Chapter 3 are reviewed and answers derived from the research outcome are provided:

- **RQ1:** How does an in-process system look like which eliminates or reduces the exposure caused by awkward posture shortly after occurrence?
- **RQ2:** How can an automated image-based ergonomics assessment be realised?
- **RQ3:** How can an automated and adaptive in-process workplace design be realised on various types of actuators?

RQ1: How does an in-process system look like which eliminates or reduces the exposure caused by awkward posture shortly after occurrence?

A system which meets these requirements has to fulfil two properties: First, the system has to be able to automatically identify ergonomically critical situations. This can be achieved through applying ergonomic assessment tools on human models obtained by a joint localisation algorithm. Algorithmic optimisations and parallel computing foster the reduction of runtime such that immediate feedback can be provided. Second, such a system must be able to automatically manipulate the workplace layout. One way to achieve this is by attaching the workpiece to a holding actuator, such as a height-adjustable platform or a robot manipulator. The task of establishing the appropriate actuator parameters can be modelled as a non-linear optimisation problem, which can then be solved by the SQP algorithm.

When the human being interacts with the system, safety is a fundamental necessity. A safe interaction can be achieved via various mechanisms. For one, actuator motions demand a human confirmation in order to be executed. This guarantees that the worker is aware of the moving actuator. Next, reducing the motion space or the degrees of freedom of the actuator can avoid dangerous configurations. Finally, reducing actuator speed makes certain that even in contact scenarios, the risk of injury is significantly diminished. Additionally, speed reduction allows the human to better react to the actuator motion.

RQ2: How can an automated image-based ergonomics assessment be realised?

Automated ergonomic assessments require that the system automatically classifies postures taking place in a process. Joint angles are suitable features to train the classifier on when it can be ensured that the localisation of the joints in the image is robust. Therefore, observing the human from the side perspective and applying 2D joint localisation algorithms supply a sufficient level of granularity while being more robust than 3D approaches. Generative methods have shown to be suitable and practical for joint localisation - they do not require training data.

RQ3: How can an automated and adaptive in-process workplace design be realised on various types of actuators?

Workplace design is comprised of identifying a set of feasible design options and opting for the most appropriate one. This process can be modelled by transforming the design task into a non-linear optimisation problem, of which the solution is the new workplace layout adjustment. In contrast to a brute-force search, optimisation algorithms can be applied, leading to a significantly lower runtime. Different types of actuators need discrete modelling of the optimisation problem terms. Although the optimisation problem formulated varies, the same non-linear solver can be applied.

9.3 Limitations of this Thesis

Answers have been supplied to the self-given research questions. However, a claim cannot be made for the completeness of the developed solution. Although essentially bridging the effectiveness-flexibility trade-off, the system necessitates strict technical and non-technical conditions to operate. Additionally, open questions arose during the course of this work. This section highlights the questions and problems that could not be solved in this thesis.

9.3.1 Focus on Posture

This work focused on preventing posture-related WMSDs. However, the risk factors, repetition and force exertion, have a similar influence on the worker's health [20]. Comprehensive WMSD prevention requires dealing with all risk factors. Addressing the other two was not possible as there were several technical problems beyond the scope of this thesis. Firstly, measuring all types of exposure requires further sensor modalities. Automatically estimating the force from image sensors has not been addressed in entirety as of yet. Secondly, different intervention schemes than the proposed method are necessary. While an actuator can reduce the force exerted by the worker, there is no intuitive answer regarding how it can decrease the physical load caused by monotonous work and repetition. Besides, it cannot be guaranteed that the developed solution will reduce postural load, as the ergonomic assessment tools it bases on are still a matter of discussion in the Human Factors area. The scientific risk described in Sec. 4.2.1 remains relevant. Nevertheless, the a solution has to be submitted by biomechanists and ergonomists.

9.3.2 3D Joint Localisation Algorithms

The realisation of the WPC is tailored to a specific camera setup. The subject is required to be captured from the side perspective, not changing the distance to the camera. The

weakest point is the joint localisation, which becomes highly challenging when required to operate in 3D. In contrast to 2D approaches, more parameters have to be simultaneously estimated. This makes the task more computationally expensive as well as harder to solve because of a higher amount of local minima. 3D real-time joint localisation algorithms have recently started to yield promising results (see Taylor et al. [121]). They have not been integrated, yet, as their implementation is challenging. Most algorithms from the new generation apply elaborated optimisation schemes and require a high amount of training data. Nevertheless, it can be expected that there will be available software libraries in the future conducive to this end.

9.3.3 Mathematical Modelling of Human Behaviour

A theoretical framework to predict human behaviour has been put forth. Although featuring promising results, this approach has two weaknesses: 1) each extension of the model requires manually formulating and adding new terms to the optimisation problem. The achievable level of complexity is limited based on the necessity to hand-craft the functional terms. A potential solution can be incorporating machine-learning methods to model complex features using a set of examples; 2) there is no theoretical guarantee concerning the quality of the solution because of the non-linear nature of the problem. A means around this could be transforming the non-linear optimisation problem into a linear or a convex one, where it can be theoretically guaranteed that the solution has a global optimum; and 3) the practicability of the approach used remains unevaluated. User studies are needed.

9.3.4 Financial Amortisation

This thesis has identified a problem and has come up with a solution from an academic perspective. The decisions made are mostly based on academic literature and studies. However, when transferring such a technology into a tangible product employed in manufacturing lines all over the world, economic aspects must be considered. The economic challenge here lies in persuading corporate decision makers that the benefits gained from such a technology surpass the costs [199]. Costs and benefits can be monetary as well as non-monetary. An example of a non-monetary benefit is employee satisfaction. Unfortunately, non-monetary costs and benefits are hard to quantify and compare with those that are monetary. Whereas costs, such as capital investment and maintenance, can be immediately noticed, the benefits are mostly non-monetary and transmit their effects after a certain period of time, e.g., around one year (according to Goggins et al. [5]). To make matters worse, first experimental results described in Sec. 8.4 imply that expensive robot actuators using the current approach do not pay off. The risk stated in Sec. 4.2.2 could not be fully resolved. Falck et al. [18] report that a way out of this dilemma requires conducting cost-benefit analyses. Another way to deal with this problem could be by introducing

governmental incentives for prevention measurements, such as lower taxes or subsidies for such technologies in order to create a quantifiable short-term benefits.

9.3.5 Ethical, Social, Legal implications (ELSI)

The WPC is a technology that requires continuous camera-based monitoring of the worker. This raises various questions concerning worker acceptance and legal implications [199]. As laws differ from country to country, this discussion is based on the case of Germany. For the WPC, one of the most relevant laws is stated in "§6b Bundesdatenschutzgesetz (BDSG)", which deals with general conditions for optical and electronic monitoring in public spaces. In brief, it states the following:

- First of all, the monitoring space has to be marked such that workers are aware of being observed by the system.
- Furthermore, if a worker can be identified by the data acquired, they have to be informed.
- When the data is no longer needed, it has to be immediately deleted.
- Finally, monitoring is only allowed when interests worth being protected ("schutzwürdige Interessen") do not dominate.

Technical solutions can help to realise the first three conditions. For instance, RGB images, which allow unambiguously identifying a worker, are not required by the WPC. Thus, only depth and silhouette images can be used for further processing. After extracting the joint angles, even all image data can be discarded. The system has to make certain that the image data does not leak outside. The final stated condition can only be achieved in a dialogue among stakeholders (management, employee board and workforce), as the interests worth being protected, the privacy threats and the benefits of the system have to be specified and gauged against one another. This discussion is most certainly outside the scope of this thesis.

9.4 Outlook

The WPC has been a specialised area of research. However, the techniques developed can benefit various other fields. Posture-related ergonomics is also a vital factor in other applications of human-robot interaction. In hand-over scenarios, it has to be guaranteed that the human can hand and receive an object in an ergonomically friendly posture. An ergonomically favourable posture is also important when humans and robots handle objects in cooperation. The advantage of the optimisation-based framework is that alternative goals can be formulated without changing the algorithm. For example, robots can facilitate the worker to apply as much force as possible or move as little as necessary

between work steps. The key value is that goals can vary while the algorithm to find the robot action remains the same.

Systems to assist ergonomic education have already been presented in Vignais et al. [9] or Lins et al. [83]. However, these solutions only notify the user that a posture is not ideal without proposing an alternative solution. The posture optimisation algorithm is able to propose to work more ergonomically, even if there is no actuator to adjust the workplace. By paying attention to the feedback, the user can learn proper working postures.

Finally, there is potential of the developed algorithms enhancing the manual workplace design process. The posture assessment component is a cost-effective way to quickly record data to be transferred into digital factory tools. The posture optimisation algorithm makes possible a digital factory tool automatically proposing promising designs that decrease the effort of workplace designers evaluating all possible options. Overall, there are prospective opportunities for this work to influence the fields of human-robot interaction as well as ergonomics.

A Posture Assessment using EAWS

The static posture section of the EAWS [34] is depicted in Fig. A.1.

The postures defined in the ergonomic assessment tools, RULA and EAWS, are characterised by the angle intervals of particular joints. Table A.1 lists the relevant angle intervals. A posture is identified by the angle intervals its joint angle values belong to. As an example, a “bent forward” posture can be independently classified from the lower limb posture (standing, sitting or kneeling) by a back angle of $20^\circ - 60^\circ$. The 3D trunk postures are explained later in this section (see Figure A.4).

Table A.1: Angle intervals of posture classes (in degrees) for joints in EAWS and REBA.

Part of the body	Class 1	Class 2	Class 3	Class 4	Min angle α_{min}	Δ_{min} (cm)
EAWS						
Back	0 – 20	20 – 60	> 60		20	18.8
Trunk (2D)	0	1 – 20	20 – 60	> 60	20	18.8
Trunk (3D)	< 10	10 – 15	15 – 25	25 – 30	5	4.8
REBA						
Upper arm	0 – 20	20 – 45	45 – 90		20	10
Lower arm	0 – 60	60 – 100	> 100		40	19
Neck	0 – 10	10 – 20	> 20		10	10
Trunk	0	1 – 20	20 – 60	> 60	20	18.8
Legs	0 – 20	20 – 60	> 60		20	10.6

Having identified the postures from the angle intervals in each frame, the EAWS score can be computed. Figure A.2 outlines an example how to calculate an ergonomic score for 2D postures using the EAWS. The occurring postures and their time share within a shift have to be determined. Each posture is then assigned an intermediate score depending on its proportion of total assessment time. In the current example, the subject bends their back without suitable support half of the time and raises their arms to shoulder level the other half. The resulting scores are 23 and 38, respectively. The final score (61 points) is calculated by summing up all intermediate scores. In order to classify a score value into one of the three risk classes, a thresholding scheme is applied. All scores below 25 points are considered as low risk processes where no action is needed. Scores from 25 to 50 points indicate a possible risk and recommend measures. Finally, if the score exceeds

50 points, the process is classified as high risk, which definitely requires action. Hence, the current example case is classified as high risk task.

Normally, the higher the EAWS, the higher the share of high-load postures. However, there are exceptions - an exemplary error case is depicted in Fig. A.3. Assume that the worker periodically performs a task where they have to adopt the posture, "elbow at shoulder level", for the majority of the time and raise their hands above their head for just a short time. The EAWS computation results of this exemplary posture sequence is written in red. Through workplace design techniques, it is possible to convert a part of the "hands above head" share into a "elbow at shoulder level" posture, which normally has a lower physical load. The EAWS results after improvements are depicted in green. As can be seen, the EAWS penalises higher proportions of a medium-load posture more than smaller proportions of a high-load posture. Hence, accepting a small occurrence of the posture "hands above head" is preferable to having a larger share of the posture, "elbow at shoulder level".

In addition to an assessment of 2D postures, the EAWS features a tool to assess asymmetry effects and 3D postures (see Fig. A.4). The load points caused by asymmetry effects are dependent on the intensity of the effect (angular value) and the duration (fraction of time spent in the posture). In total, asymmetry effects can lead to up to 30 EAWS posture points.

Ergonomic Assessment Worksheet V1.3.3																																						
Basic Positions / Postures and movements of trunk and arms (per shift)												Postures																										
(incl. loads of <3 kg and action forces of 30-40 N) Static postures: > 4sec High frequency movements: 2 trunk bending or 10 arm lifting > 60° per min												Evaluation of static postures and/or high frequent movements of trunk/arms						Sum of lines			Asymmetry effects																	
												Duration [sec/min] = $\frac{\text{duration of posture(s)} \times 60}{\text{cycle time}}$						Trunk Rotation 1) Lateral Bending 1) Far Reach 2) 			int dur			int dur			int dur											
												[%]									0-5 0-3 0-5 0-3 0-5 0-2			Intensity x Duration			Intensity x Duration			Intensity x Duration								
												[sec/min]									5 7,5 10 15 20 27 33 50 67 83			0-5 0-3 0-5 0-3 0-5 0-2			Intensity x Duration			Intensity x Duration			Intensity x Duration					
[min/8h]						24 36 48 72 96 130 160 240 320 400																																
Standing (and walking)																																						
1		Standing & walking in alternation, standing with support	0	0	0	0	0,5	1	1	1	1,5	2																										
2		Standing, no body support (for other restrict. see Extra Points)	0,7	1	1,5	2	3	4	6	8	11	13																										
3		Bent forward (20-60°) with suitable support	2	3	5	7	9,5	12	18	23	32	40																										
4		Strongly bent forward (>60°) with suitable support	3,3	5	8,5	12	17	21	30	38	51	63																										
5		Upright with elbow at / above shoulder level	3,3	5	8,5	12	17	21	30	38	51	63																										
6		Upright with hands above head level	5,3	8	14	19	26	33	47	60	80	100																										
Sitting																																						
7		Upright with back support slightly bent forward or backward	0	0	0	0	0	0,5	1	1,5	2																											
8		Upright no back support (for other restrict. see Extra Points)	0	0	0,5	1	1,5	2	3	4	5,5	7																										
9		Bent forward	0,7	1	1,5	2	3	4	6	8	11	13																										
10		Elbow at / above shoulder level	2,7	4	7	10	13	16	23	30	40	50																										
11		Hands above head level	4	6	10	14	20	25	35	45	60	75																										
Kneeling or crouching																																						
12		Upright	3,3	5	7	9	12	15	21	27	36	45																										
13		Bent forward	4	6	10	14	20	25	35	45	60	75																										
14		Elbow at / above shoulder level	6	9	16	23	33	43	62	80	108	135																										
Lying or climbing																																						
15		(Lying on back, breast or side) arms above head	6	9	15	21	29	37	53	68	91	113																										
16		Climbing	6,7	10	22	33	50	66																														
1) Trunk int 0 1 3 5 slightly 15° medium 25° strongly 30° extreme 0 1,5 2,5 3 dur never 4 sec 10 sec 13 sec 0% 6% 15% 20%															2) Far Reach int 0 1 3 5 close 60% 80% arm stretched 0 1 1,5 2 dur never 4 sec 10 sec 13 sec 0% 6% 15% 20%															Σ (max. = 15)			Σ (max. = 15)			Σ (max. = 10)		
Attention: Max. duration of evaluation = duration of task or 100%!															Attention: correct evaluation, if duration of evaluation ≠ 60s																							
Postures = Σ lines 1 - 16															(a) + (b) =																							

Figure A.1: Posture section of the EAWS.

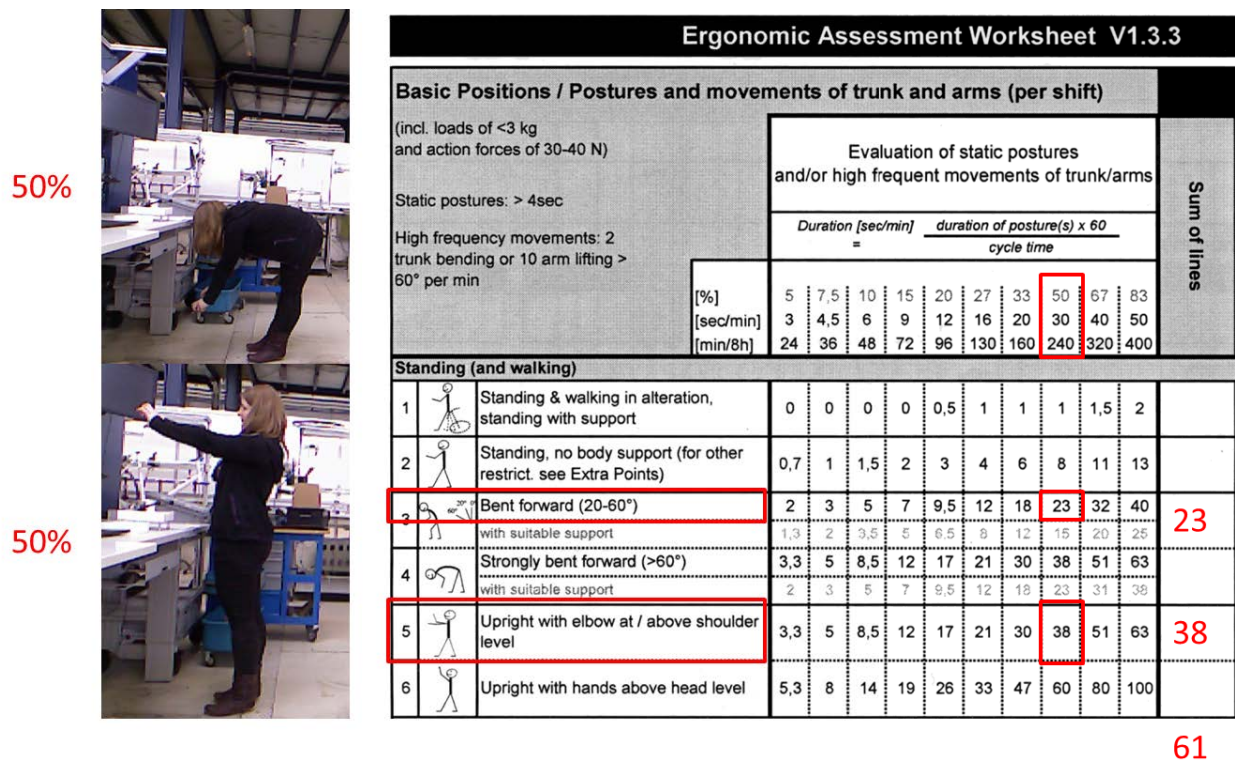


Figure A.2: Example of how to calculate an ergonomic risk score.

Ergonomic Assessment Worksheet V1.3.3																							
Basic Positions / Postures and movements of trunk and arms (per shift)													Sum of lines										
(incl. loads of <3 kg and action forces of 30-40 N) Static postures: > 4sec High frequency movements: 2 trunk bending or 10 arm lifting > 60° per min																							
														Evaluation of static postures and/or high frequent movements of trunk/arms									
														Duration [sec/min] $\frac{\text{duration of posture(s)} \times 60}{\text{cycle time}}$									
														[%] [sec/min] [min/8h]									
Standing (and walking)																							
1		Standing & walking in alternation, standing with support	0	0	0	0	0,5	1	1	1	1,5	2											
2		Standing, no body support (for other restrict. see Extra Points)	0,7	1	1,5	2	3	4	6	8	11	13											
3		Bent forward (20-60°)	2	3	5	7	9,5	12	18	23	32	40											
		with suitable support	1,3	2	3,5	5	6,5	8	12	15	20	25											
4		Strongly bent forward (>60°)	3,3	5	8,5	12	17	21	30	38	51	63											
		with suitable support	2	3	5	7	9,5	12	18	23	31	38											
5		Upright with elbow at / above shoulder level	3,3	5	8,5	12	17	21	30	38	51	63	51										
6		Upright with hands above head level	5,3	8	14	19	26	33	47	60	80	100	26										

80% 85%

20% 15%

7782

B DIN 33402: Mean Lengths of Body Parts in cm

Table B.1: Lengths of parts of the body (fifth percentile) according DIN 33402.

Part of the body	Male	Female	Min length
Upper arm	33	29	29
Lower arm	32.5	29.5	29.5
Trunk	58.5	55	55
Upper Leg	33	31	31
Lower Leg	43	40	40

C The SQP Framework

Given a non-linear program :

$$\begin{aligned} & \min_{\mathbf{x}} f(\mathbf{x}) \\ & \text{subject to} \\ & \quad g_i(\mathbf{x}) = 0 \\ & \quad h_j(\mathbf{x}) \leq 0 \end{aligned} \tag{C.1}$$

where f , g_i and h_j are possibly non-linear functions. The constraints, g_i and h_j , are combined into the vector-valued functions, g and h . To reformulate a constrained problem into a non-constrained problem, the Lagrangian of problem C.1 is introduced as:

$$\mathcal{L}(\mathbf{x}, \lambda, \sigma) = f(\mathbf{x}) - \lambda^T g(\mathbf{x}) - \sigma^T h(\mathbf{x}) \tag{C.2}$$

where λ and σ are Lagrange multipliers. In iteration, k , with the current solution, \mathbf{x}_k , the SQP algorithm in its basic form approximates problem C.1 with the following quadratic optimisation problem or quadratic program (QP):

$$\begin{aligned} & \min_{\mathbf{d}} f(\mathbf{x}_k) - \nabla f(\mathbf{x}_k)^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \text{Hess} f(\mathbf{x}_k) \mathbf{d} \\ & \text{subject to} \\ & \quad g(\mathbf{x}_k) + \nabla g(\mathbf{x}_k)^T \mathbf{d} = 0 \\ & \quad h(\mathbf{x}_k) + \nabla h(\mathbf{x}_k)^T \mathbf{d} \leq 0 \end{aligned} \tag{C.3}$$

where \mathbf{d} denotes the search direction for the new iteration, \mathbf{x}_{k+1} , $\nabla f(\mathbf{x}_k)$, the gradient of f at \mathbf{x}_k and $\text{Hess} f(\mathbf{x}_k)$, the Hessian matrix of f at \mathbf{x}_k . Problem C.3 is the result of a second-order Taylor approximation of the objective, f , and a first-order approximation of the constraints, g and h . The QP C.3 has to be solved to attain the search direction, \mathbf{d} . Given the rewritten QP:

$$\begin{aligned} & \mathbf{d}^* = \min_{\mathbf{d}} \frac{1}{2} \mathbf{d}^T \mathbf{Q} \mathbf{d} + \mathbf{d}^T \mathbf{c} \\ & \text{subject to} \\ & \quad \mathbf{A} \mathbf{d} \leq \mathbf{b} \end{aligned} \tag{C.4}$$

with its Lagrangian

$$\mathcal{L}(\mathbf{d}, \lambda, \sigma) = \frac{1}{2} \mathbf{d}^T \mathbf{Q} \mathbf{d} + \mathbf{d}^T \mathbf{c} - \lambda(\mathbf{A} \mathbf{d} - \mathbf{b}) \tag{C.5}$$

To minimise the Lagrangian, its partial derivatives, $\nabla_{\mathbf{d}}$ and ∇_{λ} , are set to zero:

$$\begin{aligned}\nabla_{\mathbf{d}}\mathcal{L} &= \mathbf{Q}\mathbf{d} + \mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda} = 0 \\ \nabla_{\lambda}\mathcal{L} &= \mathbf{A}\mathbf{d} + \mathbf{b} = 0\end{aligned}\tag{C.6}$$

Rewriting eq. C.6 yields the system of linear equations:

$$\begin{bmatrix} \mathbf{Q} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \lambda \end{bmatrix} = \begin{bmatrix} -\mathbf{c} \\ \mathbf{b} \end{bmatrix}\tag{C.7}$$

Having determined the search direction, \mathbf{d} , line search methods [193] can be employed to find the step size, α . Afterwards, the current iteration is updated with:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha\mathbf{d}\tag{C.8}$$

The final solution is reached when \mathbf{x}_k converges. The original non-linear program C.1 possibly contains multiple local minima. Which of the minima the algorithm arrives at depends on the implementation of the framework used and on the initialisation, \mathbf{x}_0 .

D Implementation Details

D.0.1 Libraries and Tools

The demonstrator system was programmed in C++ to meet the performance requirements. Moreover, the external libraries for the WPC software necessitated a C++ interface. The code was implemented on a Windows 10 (Microsoft Corporation, Redmond, WA, USA) system, however, porting the software to other operating systems is possible as platform-independent code and libraries were employed. In Windows, the Visual Studio 12 (Microsoft Corporation, Redmond, WA, USA) compiler was used. Similarly, the g++ compiler could be applied in Linux.

CMake

An important tool to achieve platform independence is CMake. CMake sets up the development environment and automatically configures paths to development tools and dependent libraries. The code uses CMake configuration files (dubbed CMakeLists.txt) to find the required libraries and development tools. CMake then utilises this information to create system-dependent development environments. Hence, only abstract configuration commands had to be written whereas the tool formulated the system specific commands.

OpenCV

OpenCV (Open Source Computer Vision Library) [200] is a library offering runtime-optimised implementations of computer vision and machine learning algorithms. Along with the algorithms, OpenCV serves up efficient data structures for images and matrices. The "cv::Mat" data structure was extensively used to exchange data between the algorithms. Moreover, the OpenCV morphological operator implementations were employed in the pre-processing step of the joint localisation algorithm.

OpenNI 2

To acquire images from the Kinect sensor, OpenNI2 [170], an open-source library to standardise an interface for natural user interfaces, was made use of. Apart from routines to acquire the image data, OpenNI2 features the proprietary middleware NiTE2 [169] to perform joint localisation. Note that only the person segmentation routine from the NiTE2 middleware was leveraged.

NLOpt

A SQP optimiser from the library, NLOpt [201, 202], was applied to solve the posture optimisation problem. Apart from the SQP implementation, NLOpt is endowed with a variety of other optimisation algorithms for various sorts of problems.

Doxygen

Doxygen was used for creating the API documentations. Doxygen enabled directly incorporating descriptions of each function into the code and acted as a tool to create class diagrams and header file descriptions from the source code.

Googletest

Exemplary unit tests were implemented in order to provide developers additional documentation and an aid when altering the code. Googletest (Google Inc., Mountain View, CA, USA) was the test framework used.

D.0.2 Division of Components into Packages

A package is a file directory containing the source, binaries, dependent libraries and test data. A user can integrate packages into their own code with CMake. The package concept can be applied in arbitrary types of projects where an exchange of different technology modules is desired.

In order to minimise the possibility that the same code was present in different packages, a packages has the ability to depend on one another. A package has a standardised directory structure where each subdirectory serves a defined purpose. Hence, a user that must use an unseen package knows where to start. Furthermore, the directory structure convention also acts as a specification for how to create new packages. A template package was incorporated such that creators of new packages can use it as a basis. In the following, the directories with their functions are listed:

- **build:** The build folder stores intermediate files created through the compilation process. Hence, the other folders are not polluted with temporary files.
- **doc:** The doc folder contains a short package documentation and API documentation created with Doxygen.
- **ext:** The ext folder features binaries of external libraries the package depends on along with test datasets.
- **install:** After compilation, the binaries, the interface definitions, dependent libraries from the ext folder and the test data are automatically copied into the install folder. This folder houses the deployable software.

- `src`: The `src` folder is where the source code is found. This includes files for unit tests located in the sub-folder, "test".

In addition to these folders, each package has configuration scripts based on CMake. These scripts configure compilers depending on the operating system and the compiler version used. Additionally, the scripts initiate copy operations of the built binaries and test data into the install folder. When further packages are needed, the configuration posts a request to the user to input their location. Furthermore, unit tests were implemented. The tests serve as an additional documentation by providing examples of how to use the API. Further, a user can monitor the feature extension process by testing the code after each feature is added. When the user introduces a bug into the existing code, particular tests are expected to fail. The testing concept requires that the user who extends the package also writes tests for the new features.

D.0.3 Performance Optimisations through Parallel Computing

For the purposes of achieving immediate feedback from the WPC algorithms, all components had to efficiently process their inputs. This turned out to be especially challenging for the joint localisation algorithms. Generative localisation requires creating and evaluating hundreds of candidate solutions in each frame. Fortunately, these operations are similar and only vary with respect to the data they are applied to. This class of problems can be well parallelised with single-instruction multiple data (SIMD) processors. Specifically, parallelisation techniques were utilised in two cases:

Multi-Core Programming using OpenMP

In the first case, the candidate search of the 2D joint localisation algorithm was parallelised using the OpenMP API [203]. OpenMP possesses functions for parallel processing on multiple processor architectures. The threads communicate across shared memory. One main drawback of OpenMP is that the communication overhead between the threads is time-intensive. This often results in worsened performance of the multi-threaded version in comparison to the single-threaded version. It is advised to perform a large partition of work rather than multiple small partitions within the threads. In the A* search routine, the first level of the search tree was divided into separate search trees, which were then processed in parallel by each thread (see Figure D.1). Afterwards, the solution with the lowest cost among each tree was computed by a single core. Starting the threads in an early phase of the search supported a speed-up of around three for an angular resolution of 10 samples with each joint.

GPU Parallelisation using the Nvidia CUDA Framework

In the second application case, the cost function for the 3D joint localisation was parallelised with the CUDA framework [204] (Nvidia Corp., Santa Clara, CA, USA). Comput-

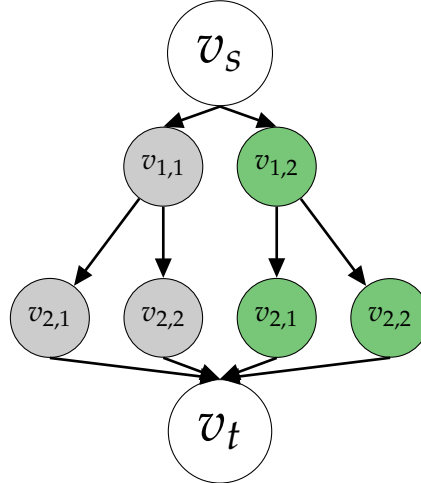
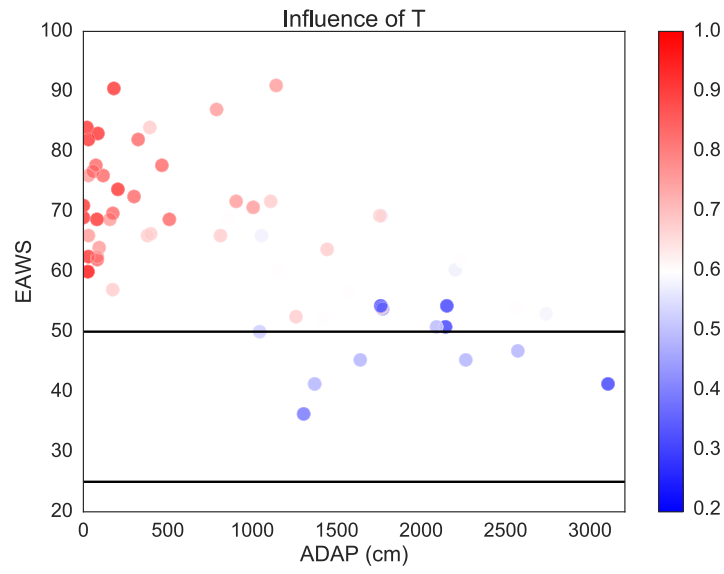


Figure D.1: Parallelisation of the tree search. One CPU core processes the green nodes while the other processes the gray nodes.

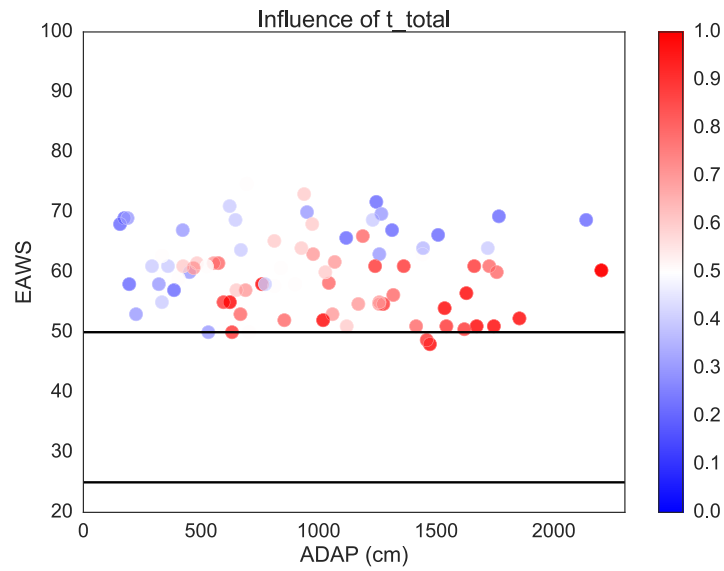
ing the cost function commences with the generation of the artificial depth image through ray-casting. For each ray, the collision point has to be computed by solving Eq. 5.12. Based on the fact that the computation of each ray is comprised of a sequence of simple operations, this solving step was parallelised on GPU hardware. In comparison to the CPU, the GPU consisted of simpler cores with low computational capacity in their own right. However, their amount was around two to three magnitudes higher and creating threads for each processing unit took significantly less time than on a CPU core. Still, there were two bottlenecks:

The first bottleneck was data transfer overhead. As the data was processed on the GPU, it had to be transferred to the GPU memory and back. Even though the GPU bandwidth, with hundreds of gigabytes per second, was sufficiently high, a call of the transfer routine adds overhead and latency, which surpasses the latency induced by the transfer. Therefore, data was to be transferred using as few partitions as possible. The second issue was the communication between the CUDA threads. Whereas the ray-object intersection computation was able to be performed independently for each thread, accumulating the values into a mean value required synchronisation between threads. The accumulation operation is called reduce operation, which is critical because multiple threads access the same memory address block. As one thread needs to wait for the intermediate result of the other, the algorithm cannot profit as strongly from the parallel processing as it does in the intersection computation step. Optimised reduce routines from CUDA's standard library, thrust [205] (Nvidia Corp., Santa Clara, CA, USA), were used. With parallelisation on GPU, the computation times of the 3D joint localisation algorithm could be decreased from a few minutes on a CPU to a few seconds per image. Although the GPU had nearly hundreds of cores, the speed-up was roughly 50 times based on the costly data transfer and reduce operations.

E Supplementary Experimental Results

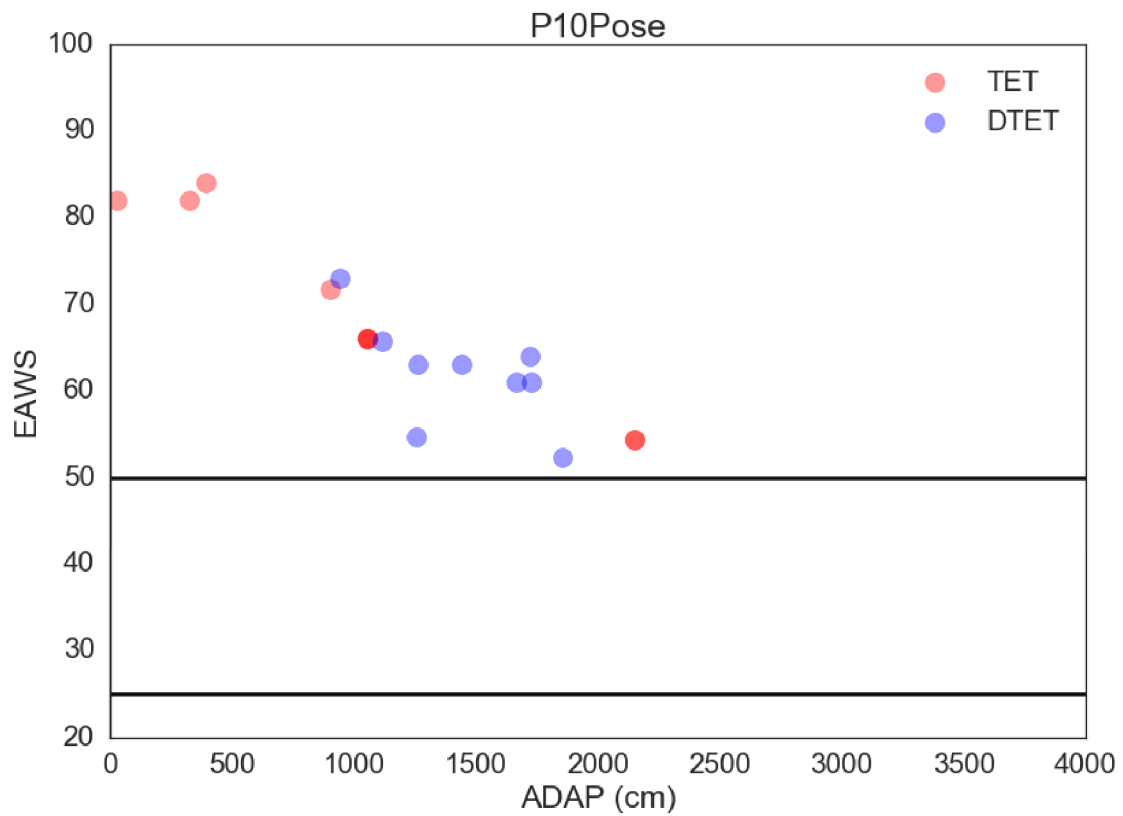


(a) TET (T)

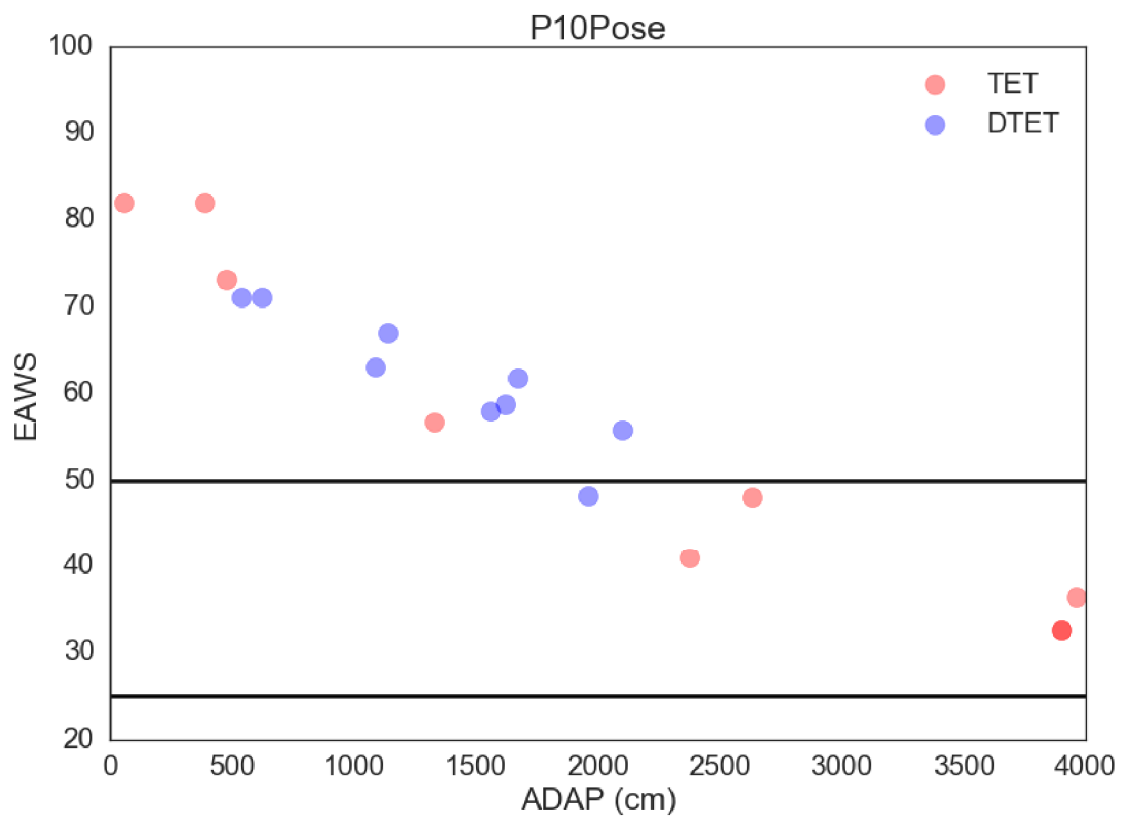


(b) DTET (t_{Total})

Figure E.1: Influence of strategy parameters on EAWS and ADAP.

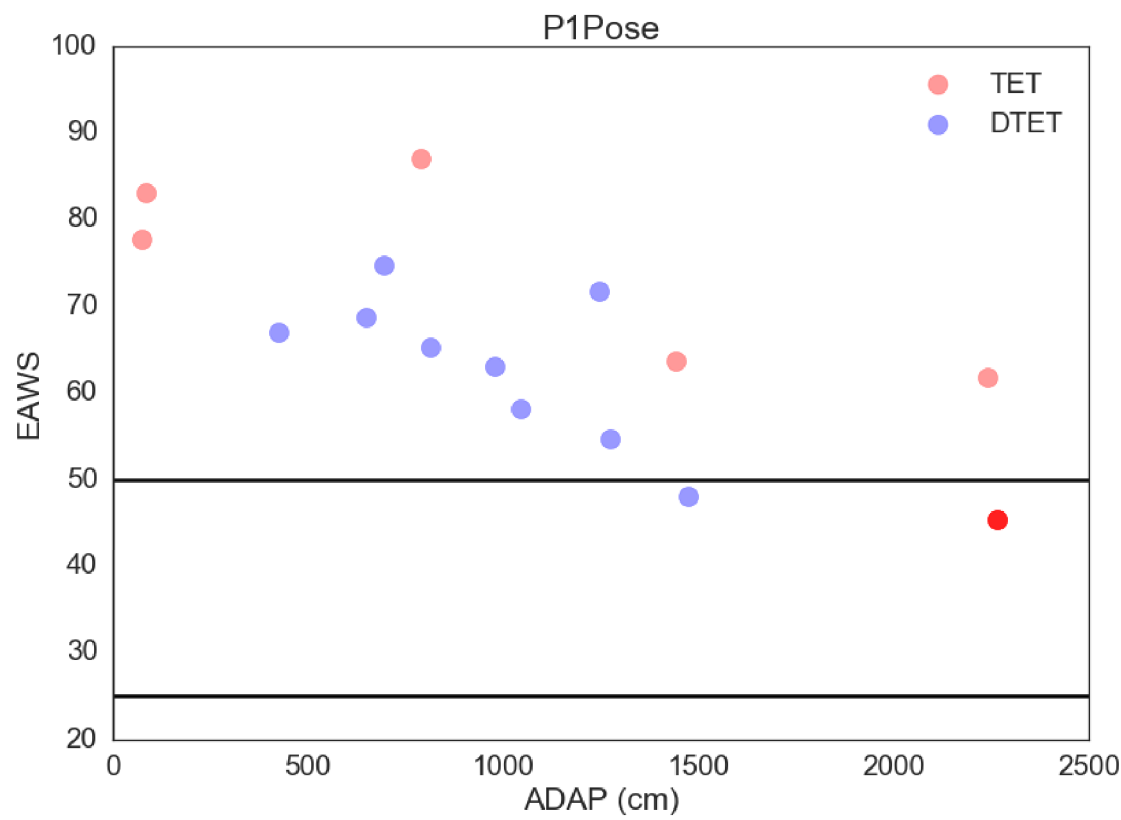


(a) Short actuator range

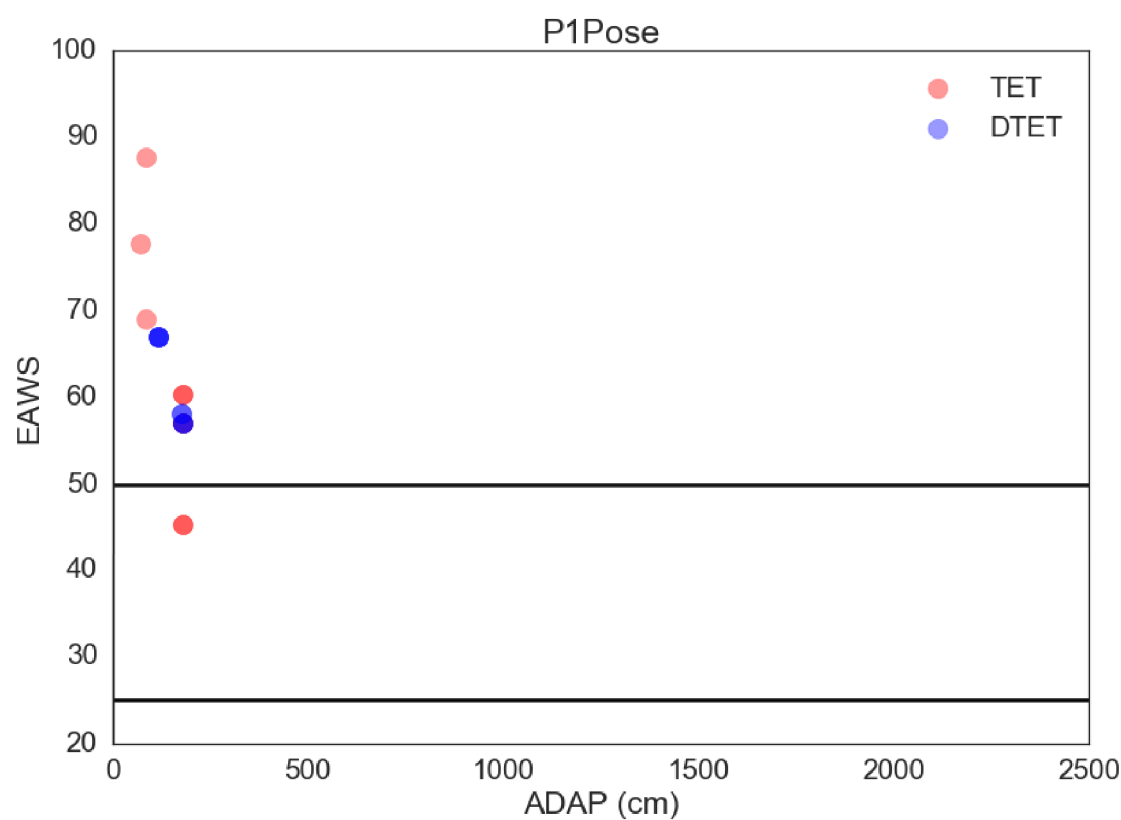


(b) Long actuator range

Figure E.2: Influence of actuator range on EAWS and ADAP.



(a) Short posture duration



(b) Long posture duration

Figure E.3: Influence of posture duration on EAWS and ADAP.

Bibliography

- [1] B. Badura, H. Schröder, J. Klose, and K. Macco, *Fehlzeiten-Report 2010: Vielfalt Managen: Gesundheit Fördern: Potenziale Nutzen Zahlen, Daten, Analysen aus allen Branchen der Wirtschaft*. Berlin, Heidelberg: Springer-Verlag, 2010.
- [2] J. Winkel and S. E. Mathiassen, "Assessment of physical work load in epidemiologic studies: concepts, issues and operational considerations," *Ergonomics*, vol. 37, no. 6, pp. 979–988, Jun. 1994.
- [3] G. Li and P. Buckle, "Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods," *Ergonomics*, vol. 42, no. 5, pp. 674–695, May 1999.
- [4] R. Ellegast, Ed., *4. Fachgespräch Ergonomie 2010 (IFA-Report 6/2011)*. Berlin: Deutsche Gesetzliche Unfallversicherung (DGUV), 2011.
- [5] R. W. Goggins, P. Spielholz, and G. L. Nothstein, "Estimating the effectiveness of ergonomics interventions through case studies: Implications for predictive cost-benefit analysis," *Journal of Safety Research*, vol. 39, no. 3, pp. 339–344, Jan. 2008.
- [6] R. Weidner, N. Kong, and J. P. Wulfsberg, "Human Hybrid Robot: a new concept for supporting manual assembly tasks," *Production Engineering*, vol. 7, no. 6, pp. 675–684, Nov. 2013.
- [7] N. J. Delleman, C. M. Haslegrave, and D. B. Chaffin, Eds., *Working postures and movements: tools for evaluation and engineering*. Boca Raton, FL: CRC Press, 2004.
- [8] K. Abdel-Malek, Z. M. Mi, J. Yang, and K. Nebel, "Optimization-based layout design," *Applied Bionics and Biomechanics*, vol. 2, no. 3-4, pp. 187–196, Mar. 2005.
- [9] N. Vignais, M. Miezal, G. Bleser, K. Mura, D. Gorecky, and F. Marin, "Innovative system for real-time ergonomic feedback in industrial manufacturing," *Applied Ergonomics*, vol. 44, no. 4, pp. 566–574, Jul. 2013.
- [10] C. Thomas, F. Busch, B. Kuhlenkoetter, and J. Deuse, "Process and Human Safety in Human-robot-interaction - a Hybrid Assistance System for Welding Applications," in *Proc. 4th International Conference on Intelligent Robotics and Applications (ICIRA) - Volume Part I*, 2011.
- [11] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-Time Human Pose Recognition in Parts from a Single Depth Image," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

- IEEE, Jun. 2011, pp. 1297 – 1304.
- [12] G. Pons-Moll and B. Rosenhahn, “Model Based pose estimation,” Jun. 2011, Tutorial at IEEE International Conference on Computer Vision (ICCV) 2011 - Looking at People: The past, the present and the future.
- [13] J. Mainprice, E. A. Sisbot, L. Jaillet, J. Cortés, R. Alami, and T. Siméon, “Planning human-aware motions using a sampling-based costmap planner,” in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 5012–5017.
- [14] E. Schneider, X. Irastorza, and S. Copsey, *OSH in figures: Work-related musculoskeletal disorders in the EU - Facts and figures*. Luxembourg: Office for Official Publ. of the Europ. Communities, 2010.
- [15] L. Punnett and D. H. Wegman, “Work-related musculoskeletal disorders: the epidemiologic evidence and the debate,” *Journal of Electromyography and Kinesiology*, vol. 14, no. 1, pp. 13–23, Feb. 2004.
- [16] A. Parent-Thirion, E. Fernández Macías, J. Hurley, and G. Vermeylen, Eds., *Fourth European working conditions survey*, ser. EF. Luxembourg: Office for Official Publ. of the European Communities, 2007.
- [17] S. Bevan, “Economic impact of musculoskeletal disorders (MSDs) on work in Europe,” *Best Practice & Research Clinical Rheumatology*, vol. 29, no. 3, pp. 356–373, Jun. 2015.
- [18] A.-C. Falck, R. Örtengren, and D. Högberg, “The impact of poor assembly ergonomics on product quality: A cost-benefit analysis in car manufacturing: Poor Assembly Ergonomics and Product Quality,” *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 20, no. 1, pp. 24–41, Jan. 2010.
- [19] I. Ben-Gal and J. Bukchin, “The ergonomic design of workstations using virtual manufacturing and response surface methodology,” *IIE Transactions*, vol. 34, no. 4, pp. 375–391, 2002.
- [20] V. Putz-Anderson, B. P. Bernard, S. E. Burt, L. L. Cole, C. Fairfield-Estill, L. J. Fine, K. A. Grant, C. Gjessing, L. Jenkins, J. J. Hurrell Jr, and others, *Musculoskeletal Disorders and Workplace Factors*. Cincinnati, OH: National Institute for Occupational Safety and Health (NIOSH), 1997.
- [21] J. M. Muggleton, R. Allen, and P. H. Chappell, “Hand and arm injuries associated with repetitive manual work in industry: a review of disorders, risk factors and preventive measures,” *Ergonomics*, vol. 42, no. 5, pp. 714–739, May 1999.
- [22] G. C. David, “Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders,” *Occupational Medicine*, vol. 55, no. 3, pp. 190–199, Mar. 2005.
- [23] P. G. Dempsey, R. W. McGorry, and W. S. Maynard, “A survey of tools and methods used by certified professional ergonomists,” *Applied Ergonomics*, vol. 36, no. 4, pp.

489–503, Jul. 2005.

- [24] W. Marras, F. Fathallah, R. Miller, S. Davis, and G. Mirka, "Accuracy of a three-dimensional lumbar motion monitor for recording dynamic trunk motion characteristics," *International Journal of Industrial Ergonomics*, vol. 9, no. 1, pp. 75–87, Jan. 1992.
- [25] R. Ellegast, I. Hermanns, and C. Schiefer, "Workload assessment in field using the ambulatory CUELA system," in *Proc. 2nd International Conference Digital Human Modeling (ICDHM)*. Springer, 2009, pp. 221–226.
- [26] M. Gazzoni, "Multichannel surface electromyography in ergonomics: Potentialities and limits," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 20, no. 4, pp. 255–271, Jul. 2010.
- [27] R. Radwin, G. Masters, and F. Lupton, "A linear force-summing hand dynamometer independent of point of application," *Applied Ergonomics*, vol. 22, no. 5, pp. 339–345, Oct. 1991.
- [28] L. Mcatamney and E. N. Corlett, "Ergonomic workplace assessment in a health care context," *Ergonomics*, vol. 35, no. 9, pp. 965–978, Sep. 1992.
- [29] S. Hignett and L. McAtamney, "Rapid Entire Body Assessment (REBA)," *Applied Ergonomics*, vol. 31, no. 2, pp. 201–205, Apr. 2000.
- [30] O. Karhu, R. Härkönen, P. Sorvali, and P. Vepsäläinen, "Observing working postures in industry: Examples of OWAS application," *Applied Ergonomics*, vol. 12, no. 1, pp. 13–17, Mar. 1981.
- [31] E. Occhipinti, "OCRA: a concise index for the assessment of exposure to repetitive movements of the upper limbs," *Ergonomics*, vol. 41, no. 9, pp. 1290–1311, Sep. 1998.
- [32] B. Buchholz, V. Paquet, L. Punnett, D. Lee, and S. Moir, "PATH: A work sampling-based approach to ergonomic job analysis for construction and other non-repetitive work," *Applied Ergonomics*, vol. 27, no. 3, pp. 177 – 187, Jul. 1996.
- [33] D. Kee and W. Karwowski, "LUBA: an assessment technique for postural loading on the upper body based on joint motion discomfort and maximum holding time," *Applied Ergonomics*, vol. 32, no. 4, pp. 357–366, 2001.
- [34] K. Schaub, G. Caragnano, B. Britzke, and R. Bruder, "The European Assembly Worksheet," *Theoretical Issues in Ergonomics Science*, vol. 14, no. 6, pp. 616–639, Nov. 2013.
- [35] E. N. Corlett, S. J. Madeley, and I. Manenica, "Posture Targeting: A Technique for Recording Working Postures," *Ergonomics*, vol. 22, no. 3, pp. 357–366, Mar. 1979.
- [36] T. R. Waters, V. Putz-Anderson, and A. Garg, *Applications manual for the revised NIOSH lifting equation*. Cincinnati, OH: U.S. Department of Health and Human Services, 1994.

- [37] I. L. Janowitz, M. Gillen, G. Ryan, D. Rempel, L. Trupin, L. Swig, K. Mullen, R. Rugulies, and P. D. Blanc, "Measuring the physical demands of work in hospital settings: Design and implementation of an ergonomics assessment," *Applied Ergonomics*, vol. 37, no. 5, pp. 641–658, Sep. 2006.
- [38] H. F. van der Molen, J. K. Sluiter, C. T. Hulshof, P. Vink, and M. H. Frings-Dresen, "Effectiveness of measures and implementation strategies in reducing physical work demands due to manual handling at work," *Scandinavian journal of work, environment & health*, vol. 2005, no. 31, pp. Suppl 2: 75–87, 2005.
- [39] V. C. Hoe, D. M. Urquhart, H. L. Kelsall, and M. R. Sim, "Ergonomic design and training for preventing work-related musculoskeletal disorders of the upper limb and neck in adults," *Cochrane Database of Systematic Reviews*, vol. 2012, no. 8, 2012.
- [40] R. Bergamasco, C. Girola, and D. Colombini, "Guidelines for designing jobs featuring repetitive tasks," *Ergonomics*, vol. 41, no. 9, pp. 1364–1383, Sep. 1998.
- [41] S. Tuncel, A. Genaidy, R. Shell, S. Salem, W. Karwowski, M. Darwish, F. Noel, and D. Singh, "Research to practice: Effectiveness of controlled workplace interventions to reduce musculoskeletal disorders in the manufacturing environment—critical appraisal and meta-analysis," *Human Factors and Ergonomics in Manufacturing*, vol. 18, no. 2, pp. 93–124, Mar. 2008.
- [42] E. H. Bos, B. Krol, A. Van Der Star, and J. W. Groothoff, "The effects of occupational interventions on reduction of musculoskeletal symptoms in the nursing profession," *Ergonomics*, vol. 49, no. 7, pp. 706–723, 2006.
- [43] S. H. van Oostrom, M. T. Driessen, H. C. de Vet, R.-L. Franche, E. Schonstein, P. Loisel, W. van Mechelen, and J. R. Anema, "Workplace interventions for preventing work disability," *Cochrane Database of Systematic Reviews*, vol. 2009, no. 2, 2009.
- [44] K. Schaub, J. Wakula, K. Berg, B. Kaiser, R. Bruder, U. Glitsch, and R.-P. Ellegast, "The Assembly Specific Force Atlas: The Assembly Specific Force Atlas," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 25, no. 3, pp. 329–339, May 2015.
- [45] J. Schmidtler, C. Hölzel, V. Knott, and K. Bengler, "Human Centered Assistance Applications for Production," *Advances in The Ergonomics in Manufacturing: Managing the Enterprise of the Future*, vol. 13, p. 380, 2014.
- [46] W. S. Marras, W. G. Allread, D. L. Burr, and F. A. Fathallah, "Prospective validation of a low-back disorder risk model and assessment of ergonomic interventions associated with manual materials handling tasks," *Ergonomics*, vol. 43, no. 11, pp. 1866–1886, 2000.
- [47] J. A. Engels, J. W. J. Van der Gulden, T. F. Senden, J. J. Kolk, and R. A. Binkhorst, "The effects of an ergonomic-educational course," *International archives of occupational and environmental health*, vol. 71, no. 5, pp. 336–342, 1998.

- [48] M. T. Driessen, K. I. Proper, M. W. van Tulder, J. R. Anema, P. M. Bongers, and A. J. van der Beek, "The effectiveness of physical and organisational ergonomic interventions on low back pain and neck pain: a systematic review," *Occupational and Environmental Medicine*, vol. 67, no. 4, pp. 277–285, Apr. 2010.
- [49] B. Das and R. M. Grady, "Industrial workplace layout design An application of engineering anthropometry," *Ergonomics*, vol. 26, no. 5, pp. 433–447, May 1983.
- [50] W. M. Keyserling, M. Brouwer, and B. Silverstein, "The effectiveness of a joint labor-management program in controlling awkward postures of the trunk, neck, and shoulders: Results of a field study," *International Journal of Industrial Ergonomics*, vol. 11, pp. 51 – 65, 1993.
- [51] M. Neuhaus, G. N. Healy, D. W. Dunstan, N. Owen, and E. G. Eakin, "Work-place Sitting and Height-Adjustable Workstations," *American Journal of Preventive Medicine*, vol. 46, no. 1, pp. 30–40, Jan. 2014.
- [52] J. Rasmussen, M. Damsgaard, S. Christensen, and E. Surma, "Design optimization with respect to ergonomic properties," *Structural and Multidisciplinary Optimization*, vol. 24, no. 2, pp. 89–97, Sep. 2002.
- [53] J. Rasmussen, J. Dahlquist, M. Damsgaard, M. de Zee, and S. T. Christensen, "Musculoskeletal modeling as an ergonomic design method," in *International Ergonomics Association XVth Triennial Conference*, 2003, pp. 24–29.
- [54] J. Rasmussen, M. Damsgaard, E. Surma, S. T. Christensen, M. de Zee, and V. Vondrak, "Anybody-a software system for ergonomic optimization," in *Fifth World Congress on Structural and Multidisciplinary Optimization*, vol. 4, 2003.
- [55] K. Abdel-Malek, J. Yang, J. H. Kim, T. Marler, S. Beck, C. Swan, L. Frey-Law, A. Mathai, C. Murphy, S. Rahmatallah, and others, "Development of the virtual-human SantosTM," in *Digital Human Modeling*, V. G. Duffy, Ed. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 490–499.
- [56] P. van der Meulen and A. Seidl, "Ramsis—the leading cad tool for ergonomic analysis of vehicles," in *Digital Human Modeling*, V. G. Duffy, Ed. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 1008–1017.
- [57] N. I. Badler, C. B. Phillips, and B. L. Webber, *Simulating humans: Computer graphics, animation, and control*. Oxford: Oxford University Press, 1993.
- [58] L. Fritzsche, R. Jendrusch, W. Leidholdt, S. Bauer, T. Jäckel, and A. Pirger, "Introducing ema (Editor for Manual Work Activities)—A New Tool for Enhancing Accuracy and Efficiency of Human Simulations in Digital Production Planning," in *Digital Human Modeling*, V. G. Duffy, Ed. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 272–281.
- [59] D. Lämkuhl, L. Hanson, and Roland Örtengren, "A comparative study of digital human modelling simulation results and their outcomes in reality: A case study within manual assembly of automobiles," *International Journal of Industrial*

- Ergonomics*, vol. 39, no. 2, pp. 428–441, Mar. 2009.
- [60] A. Enomoto, N. Yamamoto, and T. Suzuki, “Automatic estimation of the ergonomics parameters of assembly operations,” *CIRP Annals - Manufacturing Technology*, vol. 62, no. 1, pp. 13–16, Jan. 2013.
- [61] F. Caputo, G. Di Gironimo, and A. Marzano, “Ergonomic optimization of a manufacturing system work cell in a virtual environment,” *Acta Polytechnica*, vol. 46, no. 5, 2006.
- [62] L. Fritzsche, “Ergonomics risk assessment with digital human models in car assembly: Simulation versus real life,” *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 20, no. 4, pp. 287–299, Jul. 2010.
- [63] J. W. Niu, X. W. Zhang, X. Zhang, and L. H. Ran, “Investigation of ergonomics in automotive assembly line using Jack,” in *Proc. IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2010, pp. 1381–1385.
- [64] A. Marzano, K. Agyapong-Kodua, and S. Ratchev, “Virtual Ergonomics and Time Optimization of a Railway Coach Assembly Line,” *Procedia CIRP*, vol. 3, pp. 555–560, Jan. 2012.
- [65] P. Maurice, V. Padois, Y. Measson, and P. Bidaud, “Experimental assessment of the quality of ergonomic indicators for collaborative robotics computed using a digital human model,” in *International Digital Human Modeling Symposium (DHM2014)*, 2014.
- [66] D. T. Pham and H. H. Onder, “A knowledge-based system for optimizing workplace layouts using a genetic algorithm,” *Ergonomics*, vol. 35, no. 12, pp. 1479–1487, Dec. 1992.
- [67] K. Abdel-Malek, W. Yu, J. Yang, and K. Nebel, “A mathematical method for ergonomic-based design: placement,” *International Journal of Industrial Ergonomics*, vol. 34, no. 5, pp. 375–394, Nov. 2004.
- [68] G. F. Rabideau and R. H. Luk, “A Monte Carlo Algorithm for Workplace Optimization and Layout Planning - Wolap,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 19, no. 2, pp. 187–192, Oct. 1975.
- [69] T. Härtel, A. Keil, A. Hoffmeyer, and B. Toledo Munoz, “Capturing and Assessment of Human Motion during Manual Assembly Operation,” in *First International Symposium on Digital Human Modeling, Lyon, France*, 2011.
- [70] Institut für Mechatronik, “alaska/Dynamicus,” <https://www.tu-chemnitz.de/ifm/produkte-html/alaskaDYNAMICUS.html>, Chemnitz, accessed: 2016-09-06.
- [71] C. P. Daphalapurkar, W. Zhu, M. C. Leu, X. F. Liu, J. K. Gilpin-Mcminn, P. H. Wu, S. D. Snodgrass, and others, “Motion Capture for Human-Centered Simulation Using Kinects,” in *ASME 2013 International Design Engineering Technical Conferences and*

- Computers and Information in Engineering Conference*. American Society of Mechanical Engineers, 2013.
- [72] M. Manns, S. Mengel, and M. Mauer, "Experimental Effort of Data Driven Human Motion Simulation in Automotive Assembly," *Procedia CIRP*, vol. 44, pp. 114–119, 2016.
- [73] P. Mårdberg, J. S. Carlson, R. Bohlin, N. Delfs, S. Gustafsson, and L. Hanson, "Using a Formal High-level Language to Instruct Manikins to Assemble Cables," *Procedia CIRP*, vol. 23, pp. 29–34, 2014.
- [74] K. Nishiwaki, M. Kuga, S. Kagami, M. Inaba, and H. Inoue, "Whole-body cooperative balanced motion generation for reaching," *International Journal of Humanoid Robotics*, vol. 2, no. 04, pp. 437–457, 2005.
- [75] C.-F. Kuo and M.-J. Wang, "Motion generation from MTM semantics," *Computers in Industry*, vol. 60, no. 5, pp. 339–348, Jun. 2009.
- [76] D. Gläser, L. Fritzsche, S. Bauer, and W. Leidholdt, "The Quest to Validate Human Motion for Digital Ergonomic Assessment – Biomechanical Studies to Improve the Human-Like Behavior of the Human Model "EMA", in *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*, Krakow, Poland, 2014, pp. 19 – 26.
- [77] N. Miyata, K. Oguri, J. Ota, and T. Arai, "Human lift-up motion generation based on identification of time-variant performance index," in *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, vol. 3. IEEE, 2002, pp. 2503–2508.
- [78] V. Hue, P. Chiron, and J. Y. Fourquet, "On automatic generation of human motion for ergonomic evaluation of workplace and tasks," in *3rd Innovative Production Machines and Systems Virtual Conference, IPROMS*, vol. 2007, 2007.
- [79] J.-Y. Fourquet, V. Hue, and P. Chiron, "Olarge: on kinematic schemes and regularization for automatic generation of human motion and ergonomic evaluation of workplaces," in *Proc. 33rd Annual Conference of the IEEE on Industrial Electronics Society (IECON)*. IEEE, 2007, pp. 2835–2840.
- [80] K. Abdel-Malek, W. Yu, and M. Jaber, "Realistic posture prediction for maximum dexterity," *SAE Technical Paper*, no. 2001-01-2110, 2001.
- [81] C. C. Martin, D. C. Burkert, K. R. Choi, N. B. Wiczorek, P. M. McGregor, R. A. Herrmann, and P. A. Beling, "A real-time ergonomic monitoring system using the Microsoft Kinect," in *IEEE Systems and Information Design Symposium (SIIDS)*. IEEE, 2012, pp. 50–55.
- [82] Z. Ding, Z. Luo, A. Causo, I. Chen, K. Yue, S. Yeo, and K. Ling, "Inertia sensor-based guidance system for upperlimb posture correction," *Medical Engineering & Physics*, vol. 35, no. 2, pp. 269–276, Feb. 2013.
- [83] C. Lins, M. Eichelberg, L. Rölker-Denker, and A. Hein, "SIRKA: Sensoranzug zur in-

- dividuellen Rückmeldung körperlicher Aktivität," in *Dokumentation der 55. Jahrestagung der DGAUM 2015*. Aachen: Geschäftsstelle der Deutschen Gesellschaft für Arbeitsmedizin und Umweltmedizin e.V., 2015, pp. 301 – 303.
- [84] Occupational Safety and Health Administration, "Guidelines for Shipyards - Ergonomics for the Prevention of Musculoskeletal Disorders," <https://www.osha.gov/dsg/guidance/shipyard-guidelines.html>, 2008, accessed: 2016-12-01.
- [85] J. Krüger, T. K. Lien, and A. Verl, "Cooperation of human and machines in assembly lines," *CIRP Annals-Manufacturing Technology*, vol. 58, no. 2, pp. 628–646, 2009.
- [86] M. Peshkin, J. E. Colgate, W. Wannasuphoprasit, C. Moore, R. B. Gillespie, P. Akella, and others, "Cobot architecture," *Robotics and Automation, IEEE Transactions on*, vol. 17, no. 4, pp. 377–390, 2001.
- [87] R. Weidner and J. Wulfsberg, "Concept and Exemplary Realization of Human Hybrid Robot for Supporting Manual Assembly Tasks," *Procedia CIRP*, vol. 23, pp. 53–58, 2014.
- [88] C.-J. Yang, J.-F. Zhang, Y. Chen, Y.-M. Dong, and Y. Zhang, "A review of exoskeleton-type systems and their key technologies," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 222, no. 8, pp. 1599–1612, Aug. 2008.
- [89] F. Busch, "Ein konzept zur abbildung des menschen in der offline-programmierung und simulation von mensch-roboter-kollaborationen," Ph.D. dissertation, Technische Universität Dortmund, 2016.
- [90] "carat robotic innovation GmbH," <http://carat-robotic.de>, accessed: 2016-09-20.
- [91] "Mystic Game Development," <http://mysticgd.com>, accessed: 2016-09-20.
- [92] J. Wolff, C. Parker, J. Borisoff, W. B. Mortenson, and J. Mattie, "A survey of stakeholder perspectives on exoskeleton technology," *Journal of neuroengineering and rehabilitation*, vol. 11, no. 1, p. 169, 2014.
- [93] J. J. Window, "The validity of using quick ergonomics assessment tools in the prediction of developing workplace musculoskeletal disorders." in *Human Factors and Ergonomic Society of Australia 42nd Annual Conference*, 2006, pp. 1–8.
- [94] R. Grüninger, E. Kuś, and R. Hüppi, "Market study on adaptive robots for flexible manufacturing systems," in *Mechatronics, 2009. ICM 2009. IEEE International Conference on*. IEEE, 2009, pp. 1–7.
- [95] H. Schmidtke and I. Jastrzebska-Fraczek, *Ergonomie: Daten zur Systemgestaltung und Begriffsbestimmungen*. München: Hanser, 2013.
- [96] S. Haddadin, A. Albu-Schaffer, and G. Hirzinger, "Requirements for Safe Robots: Measurements, Analysis and New Insights," *The International Journal of Robotics Research*, vol. 28, no. 11-12, pp. 1507–1527, Nov. 2009.

- [97] "ISO 10218-1:2011 - Robots for industrial environments – Safety requirements – Part 1: Robot," International Organization for Standardization, ISO 10218-1:2011, 2011.
- [98] "DIN EN 60204-1:2014-10 - Sicherheit von Maschinen - Elektrische Ausrüstung von Maschinen - Teil 1: Allgemeine Anforderungen (IEC 44/709/CDV:2014); Deutsche Fassung FprEN 60204-1:2014," Deutsches Institut für Normung, DIN EN 60204-1:2014-10, 2014.
- [99] "ISO/TS 15066 - Robots and robotic devices — Collaborative robots," International Organization for Standardization, ISO 15066, 2016.
- [100] P. Domingos, "A few useful things to know about machine learning," *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, 2012.
- [101] R. Cucchiara, A. Prati, and R. Vezzani, "Posture classification in a multi-camera indoor environment," in *Proc. IEEE International Conference on Image Processing (ICIP)*, vol. 1. IEEE, 2005.
- [102] C.-F. Juang and C.-M. Chang, "Human Body Posture Classification by a Neural Fuzzy Network and Home Care System Application," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 37, no. 6, pp. 984–994, Nov. 2007.
- [103] N. Zerrouki and A. Houacine, "Automatic Classification of Human Body Postures Based on the Truncated SVD," *Journal of Advances in Computer Networks*, vol. 2, no. 1, pp. 58–62, 2014.
- [104] G. Diraco, A. Leone, and P. Siciliano, "An active vision system for fall detection and posture recognition in elderly healthcare," in *Proc. IEEE Conference on Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2010, pp. 1536–1541.
- [105] T.-L. Le, M.-Q. Nguyen, and T.-T.-M. Nguyen, "Human posture recognition using human skeleton provided by Kinect," in *Proc. IEEE International Conference on Computing, Management and Telecommunications (ComManTel)*. IEEE, 2013, pp. 340–345.
- [106] S. Monir, S. Rubya, and H. S. Ferdous, "Rotation and scale invariant posture recognition using Microsoft Kinect skeletal tracking feature," in *Proc. 12th International Conference on Intelligent Systems Design and Applications (ISDA)*. IEEE, 2012, pp. 404–409.
- [107] S. Pellegrini and L. Iocchi, "Human Posture Tracking and Classification through Stereo Vision and 3d Model Matching," *EURASIP Journal on Image and Video Processing*, vol. 2008, pp. 1–12, 2008.
- [108] A. Just, Y. Rodriguez, and S. Marcel, "Hand posture classification and recognition using the modified census transform," in *Proc. 7th International Conference on Automatic Face and Gesture Recognition (FGR)*. IEEE, 2006, pp. 351–356.
- [109] A. Kuznetsova, L. Leal-Taixé, and B. Rosenhahn, "Real-time sign language recog-

- nition using a consumer depth camera," in *Proc. IEEE International Conference on Computer Vision Workshops (ICCVW)*. IEEE, 2013, pp. 83–90.
- [110] F. Buccolieri, C. Distanto, and A. Leone, "Human posture recognition using active contours and radial basis function neural network," in *Proc. IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 2005, pp. 213–218.
- [111] S. S. Hussain, H. Husain, S. A. Samad, and N. M. Tahir, "A Simplified Shock Graph for Human Posture Classification Using the Adaptive Neuro Fuzzy Inference System," *Journal of Information & Computational Science*, vol. 9, no. 8, pp. 2035–2048, 2012.
- [112] V. Escorcia, M. A. Dávila, M. Golparvar-Fard, and J. C. Niebles, "Automated vision-based recognition of construction worker actions for building interior construction operations using RGBD cameras," in *Construction Research Congress*, 2012, pp. 879–888.
- [113] J. Wang, Z. Liu, Y. Wu, and J. Yuan, "Mining actionlet ensemble for action recognition with depth cameras," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2012, pp. 1290–1297.
- [114] S. Hauberg and K. S. Pedersen, "Spatial measures between human poses for classification and understanding," in *Articulated Motion and Deformable Objects*, P. Lopez, F. Jose, R. B. Fisher, and T. B. Moeslund, Eds. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 26–36.
- [115] X. Yang and Y. Tian, "Eigenjoints-based action recognition using naive-bayes-nearest-neighbor," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2012, pp. 14–19.
- [116] X. Yang and Y. L. Tian, "Effective 3d action recognition using eigenjoints," *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 2–11, 2014.
- [117] R. Cucchiara, C. Grana, A. Prati, and R. Vezzani, "Probabilistic Posture Classification for Human-Behavior Analysis," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 35, no. 1, pp. 42–54, Jan. 2005.
- [118] A. Jalal, S. Lee, J. T. Kim, and T.-S. Kim, "Human activity recognition via the features of labeled depth body parts," in *Proc. 10th International Conference on Smart Homes and Health Telematics (ICOST)*. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 246–249.
- [119] J. Sung, C. Ponce, B. Selman, and A. Saxena, "Unstructured human activity detection from rgbd images," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2012, pp. 842–849.
- [120] L. Xia, C.-C. Chen, and J. K. Aggarwal, "View invariant human action recognition using histograms of 3d joints," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2012, pp. 20–27.

- [121] J. Taylor, B. Luff, A. Topalian, E. Wood, S. Khamis, P. Kohli, S. Izadi, R. Banks, A. Fitzgibbon, J. Shotton, L. Bordeaux, T. Cashman, B. Corish, C. Keskin, T. Sharp, E. Soto, D. Sweeney, and J. Valentin, "Efficient and precise interactive hand tracking through joint, continuous optimization of pose and correspondences," *ACM Transactions on Graphics*, vol. 35, no. 4, pp. 1–12, Jul. 2016.
- [122] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker, "Shock graphs and shape matching," *International Journal of Computer Vision*, vol. 35, no. 1, pp. 13–32, 1999.
- [123] I. Oikonomidis, N. Kyriazis, and A. A. Argyros, "Full dof tracking of a hand interacting with an object by modeling occlusions and physical constraints," in *Proc. IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2011, pp. 2088–2095.
- [124] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1. IEEE, 2005, pp. 886–893.
- [125] M. Raptis, D. Kirovski, and H. Hoppe, "Real-time classification of dance gestures from skeleton animation," in *Proc. ACM SIGGRAPH/Eurographics symposium on computer animation*. ACM, 2011, pp. 147–156.
- [126] M. Reyes, G. Domínguez, and S. Escalera, "Featureweighting in dynamic time-warping for gesture recognition in depth data," in *Proc. IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2011, pp. 1182–1188.
- [127] A. Yao, J. Gall, and L. Van Gool, "Coupled Action Recognition and Pose Estimation from Multiple Views," *International journal of computer vision*, pp. 1–22, 2012.
- [128] R. Poppe, "Vision-based human motion analysis: An overview," *Computer vision and image understanding*, vol. 108, no. 1-2, pp. 4–18, 2007.
- [129] T. B. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231–268, Mar. 2001.
- [130] X. Ji and H. Liu, "Advances in View-Invariant Human Motion Analysis: A Review," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 13–24, Jan. 2010.
- [131] J. Gall, B. Rosenhahn, T. Brox, and H.-P. Seidel, "Optimization and Filtering for Human Motion Capture," *International Journal of Computer Vision*, vol. 87, no. 1-2, pp. 75–92, Nov. 2008.
- [132] T. Sharp, Y. Wei, D. Freedman, P. Kohli, E. Krupka, A. Fitzgibbon, S. Izadi, C. Keskin, D. Robertson, J. Taylor, J. Shotton, D. Kim, C. Rhemann, I. Leichter, and A. Vinnikov, "Accurate, Robust, and Flexible Real-time Hand Tracking," in *Proc. 33rd Annual ACM Conference on Human Factors in Computing Systems*. New York City: ACM Press, 2015, pp. 3633–3642.

- [133] C. Plagemann, V. Ganapathi, D. Koller, and S. Thrun, "Real-time identification and localization of body parts from depth images," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2010, pp. 3108–3113.
- [134] C. Keskin, F. Kırac, Y. E. Kara, and L. Akarun, "Real time hand pose estimation using depth sensors," in *Consumer Depth Cameras for Computer Vision*. Berlin, Heidelberg: Springer-Verlag, 2013, pp. 119–137.
- [135] J. Shotton, P. Kohli, A. Criminisi, A. Kipman, and A. Blake, "Efficient Human Pose Estimation from Single Depth Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 12, pp. 2821–2840, 2013.
- [136] R. Girshick, J. Shotton, P. Kohli, A. Criminisi, and A. Fitzgibbon, "Efficient regression of general-activity human poses from depth images," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 415–422.
- [137] G. Fanelli, T. Weise, J. Gall, and L. Van Gool, "Real time head pose estimation from consumer depth cameras," in *Pattern Recognition - 33rd DAGM Symposium*, R. Mester and M. Felsberg, Eds. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 101–110.
- [138] P. Peursum, S. Venkatesh, and G. West, "A Study on Smoothing for Particle-Filtered 3d Human Body Tracking," *International Journal of Computer Vision*, vol. 87, no. 1-2, pp. 53–74, Jan. 2009.
- [139] I. Mikić, M. Trivedi, E. Hunter, and P. Cosman, "Human body model acquisition and tracking using voxel data," *International Journal of Computer Vision*, vol. 53, no. 3, pp. 199–223, 2003.
- [140] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of basic engineering*, vol. 83, no. 1, pp. 95–108, 1961.
- [141] L. Zhang, J. Sturm, D. Cremers, and D. Lee, "Real-time Human Motion Tracking using Multiple Depth Cameras," *Quadrant*, vol. 1, no. a1, p. a2, 2012.
- [142] J. Bandouch, F. Engstler, and M. Beetz, "Evaluation of Hierarchical Sampling Strategies in 3d Human Pose Estimation." in *Proc. 19th British Machine Vision Conference (BMVC)*, 2008, pp. 1–10.
- [143] F. Caillette, A. Galata, and T. Howard, "Real-time 3-D human body tracking using learnt models of behaviour," *Computer Vision and Image Understanding*, vol. 109, no. 2, pp. 112–125, Feb. 2008.
- [144] R. Plänkers and P. Fua, "Tracking and Modeling People in Video Sequences," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 285–302, Mar. 2001.
- [145] S. Weik and C. E. Liedtke, "Hierarchical 3d pose estimation for articulated human body models from a sequence of volume data," in *Proc. International Workshop on Robot Vision (RobVis)*, 2001, pp. 27–34.
- [146] C. Stoll, N. Hasler, J. Gall, H. Seidel, and C. Theobalt, "Fast articulated motion track-

- ing using a sums of Gaussians body model," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 951–958.
- [147] L. A. Schwarz, A. Mkhitarian, D. Mateus, and N. Navab, "Human skeleton tracking from depth data using geodesic distances and optical flow," *Image and Vision Computing*, vol. 30, no. 3, pp. 217–226, Mar. 2012.
- [148] X. Wei, P. Zhang, and J. Chai, "Accurate realtime full-body motion capture using a single depth camera," *ACM Transactions on Graphics*, vol. 31, no. 6, p. 1, Nov. 2012.
- [149] C. Qian, X. Sun, Y. Wei, X. Tang, and J. Sun, "Realtime and robust hand tracking from depth," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1106–1113.
- [150] A. Elhayek, E. de Aguiar, A. Jain, J. Tompson, L. Pishchulin, M. Andriluka, C. Bregler, B. Schiele, and C. Theobalt, "Efficient ConvNet-based marker-less motion capture in general scenes with a low number of cameras," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2015, pp. 3810–3818.
- [151] S. Li, H. Lu, X. Ruan, and Y. Chen, "Pose estimation and body segmentation based on hierarchical searching tree," in *Proc. 18th IEEE International Conference on Image Processing (ICIP)*, 2011, pp. 1289–1292.
- [152] I. Oikonomidis, N. Kyriazis, and A. A. Argyros, "Efficient model-based 3d tracking of hand articulations using Kinect." in *Proc. 22nd British Machine Vision Conference (BMVC)*, 2011.
- [153] X. S. Nguyen, S. Dubuisson, and C. Gonzales, "Hierarchical Annealed Particle Swarm Optimization for Articulated Object Tracking," in *Computer Analysis of Images and Patterns*. Berlin, Heidelberg: Springer-Verlag, 2013, pp. 319–326.
- [154] S. Saini, N. Zakaria, D. R. A. Rambli, and S. Sulaiman, "Markerless Human Motion Tracking Using Hierarchical Multi-Swarm Cooperative Particle Swarm Optimization," *PLOS ONE*, vol. 10, no. 5, p. e0127833, May 2015.
- [155] V. John, S. Ivekovic, and E. Trucco, "Articulated Human Motion Tracking with HPSO." in *Proc. 4th International Conference on Computer Vision Theory and Applications (VISAPP)*, 2009, pp. 531–538.
- [156] V. John, E. Trucco, and S. Ivekovic, "Markerless human articulated tracking using hierarchical particle swarm optimisation," *Image and Vision Computing*, vol. 28, no. 11, pp. 1530–1547, Nov. 2010.
- [157] S. Suwajanakorn, C. Hernandez, and S. M. Seitz, "Depth from Focus with Your Mobile Phone," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3497–3506.
- [158] Y. Han, J.-Y. Lee, and I. So Kweon, "High quality shape from a single rgb-d image under uncalibrated natural illumination," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2013, pp. 1617–1624.

- [159] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 1. IEEE, 2003, pp. I–195.
- [160] M. Kytö, M. Nuutinen, and P. Oittinen, "Method for measuring stereo camera depth accuracy based on stereoscopic vision," in *Proc. SPIE 7864, Three-Dimensional Imaging, Interaction, and Measurement*. International Society for Optics and Photonics, 2011, pp. 78 640I–78 640I.
- [161] Jungong Han, Ling Shao, Dong Xu, and J. Shotton, "Enhanced Computer Vision With Microsoft Kinect Sensor: A Review," *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1318–1334, Oct. 2013.
- [162] K. Khoshelham and S. O. Elberink, "Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications," *Sensors*, vol. 12, no. 12, pp. 1437–1454, Feb. 2012.
- [163] H. Sarbolandi, D. Lefloch, and A. Kolb, "Kinect range sensing: Structured-light versus Time-of-Flight Kinect," *Computer Vision and Image Understanding*, vol. 139, pp. 1–20, Oct. 2015.
- [164] Y. Schröder, A. Scholz, K. Berger, K. Ruhl, S. Guthe, and M. Magnor, "Multiple kinect studies," *Computer Graphics*, vol. 2, no. 4, p. 6, 2011.
- [165] D. A. Butler, S. Izadi, O. Hilliges, D. Molyneaux, S. Hodges, and D. Kim, "Shake'n'sense: reducing interference for overlapping structured light depth cameras," in *Proc. SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1933–1936.
- [166] A. Maimone and H. Fuchs, "Reducing interference between multiple structured light depth sensors using motion," in *IEEE Virtual Reality Short Papers and Posters (VRW)*. IEEE, 2012, pp. 51–54.
- [167] L. Yang, L. Zhang, H. Dong, A. Alelaiwi, and A. E. Saddik, "Evaluating and Improving the Depth Accuracy of Kinect for Windows v2," *IEEE Sensors Journal*, vol. 15, no. 8, pp. 4275–4285, Aug. 2015.
- [168] "DIN 33402-2:2005-12 - Ergonomie - Körpermaße des Menschen - Teil 2: Werte," Deutsches Institut für Normung, DIN 33402-2:2005-12, 2005.
- [169] "NiTE 2 middleware," <http://openni.ru/files/nite/>, accessed: 2016-09-07.
- [170] "OpenNI 2 SDK," <http://structure.io/openni>, accessed: 2016-09-07.
- [171] P. B. Laval, "Mathematics for Computer Graphics-Barycentric Coordinates," Kenesaw State University, Tech. Rep., 2003.
- [172] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.
- [173] A. Fey, "Posenschätzung für die Ergonomieanalyse aus Binärbildern mittels gra-

- dientenbasierter Optimierung,” Bachelor’s Thesis, Technische Universität Berlin, 2015.
- [174] R. C. Eberhart and Y. Shi, “Particle swarm optimization: developments, applications and resources,” in *Proc. Congress on Evolutionary Computation*. IEEE, 2001, pp. 81–86.
- [175] C. J. Bastos-Filho, J. D. Andrade, M. R. Pita, and A. D. Ramos, “Impact of the quality of random numbers generators on the performance of particle swarm optimization,” in *Proc. IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 2009, pp. 4988–4993.
- [176] S. Xu and Y. Rahmat-Samii, “Boundary Conditions in Particle Swarm Optimization Revisited,” *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 3, pp. 760–765, Mar. 2007.
- [177] D. Kieneck, “Beschleunigung des generativen Trackings durch Parallelisierung auf der Grafikkarte,” Master’s Thesis, Technische Universität Berlin, 2016.
- [178] K. Abdel-Malek, J. Yang, T. Marler, S. Beck, A. Mathai, X. Zhou, A. Patrick, and J. Arora, “Towards a new generation of virtual humans,” *International Journal of Human Factors Modelling and Simulation*, vol. 1, no. 1, pp. 2–39, 2006.
- [179] J. Yang, R. T. Marler, H. Kim, J. Arora, and K. Abdel-Malek, “Multi-objective optimization for upper body posture prediction,” in *Proc. 10th AIAA/ISSMO multidisciplinary analysis and optimization conference*, vol. 30, 2004.
- [180] R. T. Marler, J. S. Arora, J. Yang, H.-J. Kim, and K. Abdel-Malek, “Use of multi-objective optimization for digital human posture prediction,” *Engineering Optimization*, vol. 41, no. 10, pp. 925–943, Oct. 2009.
- [181] Z. Dong, J. Xu, N. Zou, and C. Chai, “Posture Prediction Based on Orthogonal Interactive Genetic Algorithm,” in *Proc. 4th International Conference on Natural Computation (ICNC)*, vol. 1. IEEE, 2008, pp. 336–340.
- [182] L. Ma, W. Zhang, D. Chablat, F. Bennis, and F. Guillaume, “Multi-objective optimisation method for posture prediction and analysis with consideration of fatigue effect and its application case,” *Computers & Industrial Engineering*, vol. 57, no. 4, pp. 1235–1246, Nov. 2009.
- [183] E. S. Jung and S. Park, “Prediction of human reach posture using a neural network for ergonomic man models,” *Computers & industrial engineering*, vol. 27, no. 1, pp. 369–372, 1994.
- [184] B. Zhang, I. Horváth, J. Molenbroek, and C. Snijders, “Using artificial neural networks for human body posture prediction,” *International Journal of Industrial Ergonomics*, vol. 40, no. 4, pp. 414–424, Jul. 2010.
- [185] H. Zhao, G. Zheng, and W. Wen, “Human-machine posture prediction and working efficiency evaluation of virtual human using radial basis function neural network,”

- in *Proc. IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS)*, vol. 1. IEEE, 2010, pp. 406–410.
- [186] M. Bataineh, T. Marler, and K. Abdel-Malek, “Artificial Neural Network-Based Prediction of Human Posture,” in *Digital Human Modeling and Applications in Health, Safety, Ergonomics, and Risk Management. Human Body Modeling and Ergonomics*, D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, and V. G. Duffy, Eds. Berlin, Heidelberg: Springer-Verlag, 2013, pp. 305–313.
- [187] A. Freivalds and B. W. Niebel, *Methods, Standards, and Work Design*. New York, NY: McGraw Hill, 2009.
- [188] X. F. Zha and S. Y. E. Lim, “Intelligent design and planning of manual assembly workstations: A neuro-fuzzy approach,” *Computers & Industrial Engineering*, vol. 44, no. 4, pp. 611–632, 2003.
- [189] W. J. Braun, R. Rebollar, and E. F. Schiller, “Computer aided planning and design of manual assembly systems,” *International Journal of Production Research*, vol. 34, no. 8, pp. 2317–2333, Aug. 1996.
- [190] S. Pavlovic-Veselinovic, A. Hedge, and M. Veselinovic, “An ergonomic expert system for risk assessment of work-related musculo-skeletal disorders,” *International Journal of Industrial Ergonomics*, vol. 53, pp. 130–139, May 2016.
- [191] S. Breznitz, *Cry wolf: The psychology of false alarms*. Hove (UK): Psychology Press, 2013.
- [192] T. Glad and L. Ljung, *Control theory*. Boca Raton, FL: CRC press, 2000.
- [193] J. Nocedal and S. J. Wright, *Numerical optimization*, 2nd ed., ser. Springer series in operations research. New York: Springer, 2006.
- [194] Universal Robots, “UR5 robot,” <https://www.universal-robots.com/products/ur5-robot/>, accessed: 2016-10-12.
- [195] S. Burbeck, “Applications programming in smalltalk-80 (tm): How to use model-view-controller (mvc),” *Smalltalk-80 v2*, vol. 5, 1992.
- [196] J. Vlissides, R. Helm, R. Johnson, and E. Gamma, “Design patterns: Elements of reusable object-oriented software,” *Reading: Addison-Wesley*, vol. 49, no. 120, p. 11, 1995.
- [197] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [198] L. v. d. Maaten and G. Hinton, “Visualizing data using t-SNE,” *Journal of Machine Learning Research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

- [199] C. Pilz, "Entwicklung eines Steuerungssystems und Anwendungsanalyse eines Ergonomie-Assistenzsystems für Montagearbeitsplätze," Master's Thesis, Technische Universität Berlin, 2014.
- [200] "Open Source Computer Vision Library," <http://opencv.org/>, accessed: 2016-10-12.
- [201] S. G. Johnson, "The NLOpt nonlinear-optimization package," <http://ab-initio.mit.edu/nlopt>, accessed: 2016-12-01.
- [202] D. Kraft, "Algorithm 733: TOMP–Fortran modules for optimal control calculations," *ACM Transactions on Mathematical Software (TOMS)*, vol. 20, no. 3, pp. 262–281, 1994.
- [203] "OpenMP API," <http://www.openmp.org/>, accessed: 2017-04-21.
- [204] "CUDA Parallel Computing Platform," <https://developer.nvidia.com/cuda-toolkit>, accessed: 2017-04-21.
- [205] "Thrust - Parallel Algorithms Library," <https://developer.nvidia.com/thrust>, accessed: 2016-11-03.