

# Bachelorarbeit

Sven Koch

Hochgenaue Positionsbestimmung mit Ultra-wideband und  
Reinforcement Learning

Sven Koch

# Hochgenaue Positionsbestimmung mit Ultra-wideband und Reinforcement Learning

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung  
im Studiengang *Bachelor of Science Informatik Technischer Systeme*  
am Department Informatik  
der Fakultät Technik und Informatik  
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer: Prof. Dr. Stephan Pareigis  
Zweitgutachter: Prof. Dr. Tim Tiedemann

Eingereicht am: 21. Dezember 2021

**Sven Koch**

**Thema der Arbeit**

Hochgenaue Positionsbestimmung mit Ultra-wideband und Reinforcement Learning

**Stichworte**

Positionsbestimmung, Ultra-wideband, Reinforcement Learning, Kalman-Filter, Particle-Filter

**Kurzzusammenfassung**

Im Rahmen dieser Bachelorarbeit wird ein Positionierungssystem vorgestellt und erprobt, das auf der Ultra-wideband Drahtlostechnologie beruht. Die Positionsbestimmung erfolgt in einem zweistufigen Verfahren mithilfe von Laufzeitmessungen. Zur Verbesserung der Positionierungsgenauigkeit wird neben einem Kalman- und Particle-Filter ein auf Reinforcement Learning basierender Ansatz eingeführt. In der experimentellen Anwendung kann gezeigt werden, dass das vorgestellte Positionierungssystem eine maximale Genauigkeit von bis zu 10 cm erreicht.

**Sven Koch**

**Title of Thesis**

High-accuracy Positioning with Ultra-wideband and Reinforcement Learning

**Keywords**

Localization, Ultra-wideband, Reinforcement Learning, Kalman Filter, Particle Filter

**Abstract**

Within the scope of this bachelor thesis, a ultra-wideband-based localization system is presented and tested. The position is estimated in a two-step manner using time of flight. To improve positioning accuracy, a reinforcement learning-based approach is introduced in addition to a classical Kalman and particle filter. The experimental validation reveals that the proposed localization system achieves a maximum accuracy of up to 10 cm.

# Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>vii</b>
<b>Tabellenverzeichnis</b>	<b>ix</b>
<b>Abkürzungen</b>	<b>x</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Ziele . . . . .	1
1.2 Vorgehensweise . . . . .	2
1.3 Struktur . . . . .	2
<b>2 Grundlagen</b>	<b>4</b>
2.1 Positionierungssysteme . . . . .	4
2.1.1 Outdoor-Positionierungssysteme . . . . .	4
2.1.2 Indoor-Positionierungssysteme . . . . .	6
2.2 Ultra-wideband . . . . .	7
2.2.1 Technologie . . . . .	7
2.2.2 Regulation . . . . .	9
2.3 Einsatzgebiete . . . . .	11
2.4 IEEE 802.15.4 . . . . .	12
2.4.1 IEEE 802.15.4 Paketaufbau . . . . .	12
2.5 Terminologie . . . . .	14
<b>3 Positionsbestimmung</b>	<b>15</b>
3.1 Signalparameter zur Positionsbestimmung . . . . .	16
3.1.1 Received Signal Strength Indicator (RSSI) . . . . .	16
3.1.2 Time of Arrival (ToA) . . . . .	17
3.1.3 Time Difference of Arrival (TDoA) . . . . .	22
3.1.4 Angle of Arrival (AoA) . . . . .	23

3.2	Methoden zur Positionsbestimmung aus den Signalparametern . . . . .	23
3.2.1	Triangulation . . . . .	23
3.2.2	Trilateration . . . . .	24
3.2.3	Hyperbelnavigation . . . . .	27
3.3	Dilution of Precision . . . . .	28
<b>4</b>	<b>Ultra-wideband Lokalisationssystem</b>	<b>30</b>
4.1	Hardwareplattform . . . . .	30
4.2	Protokoll zur Entfernungsmessung . . . . .	32
4.2.1	Blink Message . . . . .	37
4.2.2	Poll Message . . . . .	37
4.2.3	Response Message . . . . .	37
4.2.4	Final Message . . . . .	38
4.3	Technische Daten . . . . .	39
4.4	Zeitmultiplexverfahren . . . . .	39
4.5	Lokalisierungseinheit . . . . .	41
4.6	Praxistest . . . . .	41
4.7	Robustheit . . . . .	43
<b>5</b>	<b>Verbesserung der Lokalisierungsgenauigkeit</b>	<b>44</b>
5.1	Fehler . . . . .	44
5.2	Reinforcement Learning . . . . .	45
5.2.1	Beobachtungsraum . . . . .	46
5.2.2	Aktionsraum . . . . .	49
5.2.3	Belohnungsmechanismus . . . . .	50
5.2.4	Reinforcement Learning Algorithmus . . . . .	52
5.2.5	Modellarchitektur . . . . .	54
5.3	Kalman-Filter . . . . .	58
5.4	Particle-Filter . . . . .	62
5.5	Vergleich . . . . .	63
<b>6</b>	<b>Fazit</b>	<b>69</b>
6.1	Ausblick . . . . .	70
	<b>Literaturverzeichnis</b>	<b>72</b>
<b>A</b>	<b>Anhang: Schaltplan UWB-Beacon</b>	<b>75</b>

<b>Glossar</b>	<b>77</b>
<b>Selbstständigkeitserklärung</b>	<b>78</b>

# Abbildungsverzeichnis

2.1	Untere, mittlere und obere Frequenz eines UWB-Signals . . . . .	8
2.2	Frequenzverteilung verschiedener Drahtlossysteme . . . . .	10
2.3	Darstellung der IEEE 802.15.4 Paketstruktur . . . . .	13
3.1	Direkter Ansatz und zweistufiger Ansatz zur Positionsbestimmung . . . . .	15
3.2	Single-Sided Two-Way Ranging (SS-TWR) . . . . .	18
3.3	Double-Sided Two-Way Ranging (DS-TWR) mit vier Nachrichten . . . . .	20
3.4	Double-Sided Two-Way Ranging (DS-TWR) mit drei Nachrichten . . . . .	21
3.5	Geometrische Darstellung der Triangulation . . . . .	24
3.6	Geometrische Darstellung der Trilateration . . . . .	25
3.7	Geometrische Darstellung der Hyperbelnavigation . . . . .	27
3.8	Dilution of Precision . . . . .	29
4.1	UWB-Beacon . . . . .	31
4.2	Sequenzdiagramm: DS-TWR im UWB-Lokalisationssystem . . . . .	33
4.3	Allgemeines Nachrichtenformat nach IEEE 802.15.4 . . . . .	34
4.4	Kodierungen der Nachrichten zur Entfernungsmessung . . . . .	34
4.5	Zeitprofil zur gemeinsamen Nutzung des Kommunikationskanals . . . . .	40
4.6	Bildschirmfoto der zentralen Lokalisierungseinheit . . . . .	42
4.7	Praxistest des UWB-Lokalisationssystems mit einem AR-Trackingsystem im CSTI . . . . .	42
5.1	Vergleich des Datenflusses einer typischen Positionsbestimmung und einer Positionsbestimmung mit nachgeschaltetem RL-Modell . . . . .	46
5.2	POMDP-Diagramm des RL-Modells . . . . .	48
5.3	Modell des RL-Agenten . . . . .	49
5.4	Aktualisierungsprozess der RL-Modelleingaben . . . . .	49
5.5	A3C Modellarchitektur . . . . .	55
5.6	Sensitivitätsanalyse der Anzahl der im Training verwendeten Threads . . . . .	57

5.7	Ausschnitt einer geradlinigen Trajektorie . . . . .	63
5.8	Ausschnitt einer kreisförmigen Trajektorie . . . . .	63
5.9	Fehler der geradlinigen Trajektorie . . . . .	65
5.10	Akkumulierter Fehler der geradlinigen Trajektorie . . . . .	65
5.11	Fehler der kreisförmigen Trajektorie . . . . .	65
5.12	Akkumulierter Fehler der kreisförmigen Trajektorie . . . . .	65
5.13	Geradlinige Trajektorie korrigiert mit dem RL-Modell . . . . .	67
5.14	Kreisförmige Trajektorie korrigiert mit dem RL-Modell . . . . .	67
5.15	Geradlinige Trajektorie korrigiert mit dem Kalman-Filter . . . . .	67
5.16	Kreisförmige Trajektorie korrigiert mit dem Kalman-Filter . . . . .	67
5.17	Geradlinige Trajektorie korrigiert mit dem Particle-Filter . . . . .	67
5.18	Kreisförmige Trajektorie korrigiert mit dem Particle-Filter . . . . .	67

# Tabellenverzeichnis

2.1	Genauigkeitsanforderungen an verschiedene Lokalisierungsanwendungen . . .	5
3.1	Typische zeitbedingte Fehler bei der SS-TWR Signallaufzeitmessung . . .	19
3.2	Typische zeitbedingte Fehler bei der DS-TWR Signallaufzeitmessung . . .	22
4.1	Aufbau des IEEE 802.15.4 Nachrichtenformats . . . . .	35
4.2	Aufbau des IEEE 802.15.4 Frame Control Feldes . . . . .	36
4.3	Felder in der Blink Message . . . . .	37
4.4	Felder in der Poll Message . . . . .	37
4.5	Felder in der Response Message . . . . .	38
4.6	Felder in der Final Message . . . . .	38

# Abkürzungen

**A3C** Asynchronous Advantage Actor Critic.

**AoA** Angle of Arrival.

**AR** Augmented Reality.

**DOP** Dilution of Precision.

**DQN** Deep Q-Learning.

**DS-TWR** Double-Sided Two-Way Ranging.

**GNSS** Global Navigation Satellite System.

**GPS** Global Positioning System.

**IMU** Inertial Measurement Unit.

**IPS** Indoor Positioning System.

**MDP** Markov Decision Process.

**POMDP** Partially Observable Markov Decision Process.

**RL** Reinforcement Learning.

**RSSI** Received Signal Strength Indicator.

**RTT** Round Trip Time.

**SLAM** Simultaneous Localization and Mapping.

**SNR** Signal-Rausch-Verhältnis.

**SS-TWR** Single-Sided Two-Way Ranging.

**TDMA** Time Division Multiple Access.

**TDoA** Time Difference of Arrival.

**ToA** Time of Arrival.

**ToF** Time of Flight.

**UWB** Ultra-wideband.

# 1 Einleitung

In den letzten Jahren ist die Bedeutung von präzisen Ortungssystemen für Innenräume stark gestiegen. Besonders gefragt ist die Technologie zum Beispiel für Industrie- und Logistikanwendungen oder zur Indoor-Navigation in großen öffentlichen Gebäuden.

Mittlerweile existieren zahlreiche Methoden zur Innenraumlokalisierung, die auf den bekannten und verbreiteten Drahtlostechnologien WLAN oder Bluetooth basieren. Diese Systeme haben jedoch oft den Nachteil, dass keine hohen Genauigkeiten bei der Positionsbestimmung erreicht werden. [20]

Im Industrieumfeld hat sich die Ultra-wideband (UWB) Technologie für Lokalisierungsanwendungen in Innenräumen etabliert. Allmählich wird diese Technik durch große Hersteller auch für den Consumer-Markt verfügbar gemacht. [4]

Mit der Ultra-wideband Funktechnologie sind Positionsbestimmungen mit einer vergleichsweise hohen Genauigkeit, geringen Latenzzeiten und niedrigen Energieverbräuchen möglich. [20]

## 1.1 Ziele

Im Rahmen dieser Arbeit soll ein leistungsfähiges und hochgenaues Lokalisationssystem basierend auf der Drahtlostechnologie Ultra-wideband entwickelt werden. Gegenüber anderen bereits kommerziell erhältlichen UWB-Lokalisationssystemen ist das gesamte Know-how und die Entwicklung (Hardware- und Softwareentwicklung) innerhalb der HAW Hamburg angesiedelt. Die Vorteile sind eine kostengünstige und individuelle Nutzung des Lokalisationssystems für Projekte und Forschungsgruppen der HAW Hamburg.

In der *autosys* Forschungsgruppe [18] der HAW Hamburg könnte eine UWB-Positionsbestimmung beispielsweise Einsatz im Bereich der Miniaturautonomie finden. Im

Rahmen des Projektes *Test Area Intelligent Quartier Mobility* (TIQ) wäre die Verwendung zudem im Außenbereich denkbar und soll besonders die Notwendigkeit von teuren Echtzeitkinematik-Systemen vermeiden.

### 1.2 Vorgehensweise

Zunächst wird das UWB-Lokalisationssystem auf einer eigenen Hardwareplattform implementiert. Dazu werden geeignete Verfahren und Drahtlosprotokolle zur Entfernungs- und Positionsmessung herausgearbeitet und umgesetzt. Die Hardwareentwicklung ist nicht Teil dieser Arbeit und wurde bereits von Herrn Markus Kasten [14] in Zusammenhang mit der Entwicklung einer Fahrplattform für autonome Miniaturfahrzeuge beschrieben.

Mit dem Lokalisationssystem werden verschiedene Algorithmen zur Verbesserung der Lokalisierungsgenauigkeit erprobt. Besonders hervorzuheben ist hierbei ein Verfahren, das auf Grundlage eines modernen Reinforcement Learning (RL) Modells basiert.

Mithilfe von Ground Truth Informationen wird die erzielte Genauigkeit des vorgestellten Lokalisationssystems ermittelt und die Praxistauglichkeit des Systems anhand der Messdaten eingeordnet.

### 1.3 Struktur

Die Bachelorarbeit ist in sechs Kapitel untergliedert:

Die Arbeit beginnt im ersten Kapitel mit einer Einleitung in die Problemstellung. Das Ziel dieser Arbeit wird vorgestellt und die Vorgehensweise zur Erarbeitung des Problems beschrieben.

Das zweite Kapitel stellt die nötigen Grundlagen für diese Arbeit zusammen. Neben allgemeinen Informationen zu Lokalisationssystemen werden auch die wichtigsten Aspekte von Ultra-wideband Signalen vorgestellt.

Im dritten Kapitel werden verschiedene Möglichkeiten zur drahtlosen Positionsbestimmung eingeführt. Neben verschiedenen Signalparametern aus denen Positionsinformationen gewonnen werden können, werden mathematische und geometrische Methoden vorgestellt, um eine Position aus diesen Parametern zu berechnen.

Das vierte Kapitel dient im Wesentlichen zur Dokumentation des entwickelten Lokalisationssystems. Die technischen Daten und die eingesetzten Protokollabläufe werden aufgeführt und dokumentiert.

Das fünfte Kapitel stellt drei verschiedene Algorithmen zur Verbesserung der Lokalisierungsgenauigkeit vor. Zusätzlich zu zwei klassischen und etablierten Verfahren wird ebenfalls ein maschinelles Lernverfahren eingesetzt, um die Genauigkeit des Systems zu verbessern. In einem Vergleich werden die Leistungen der einzelnen Algorithmen gegenübergestellt und die Genauigkeit des entwickelten Lokalisationssystems ermittelt.

Abschließend werden die Ergebnisse dieser Arbeit im sechsten Kapitel resümiert und Ansätze für Weiterentwicklungsmöglichkeiten angeführt.

## 2 Grundlagen

In diesem Kapitel wird allgemein in die Grundlagen von Positionierungssystemen und der Ultra-wideband Technologie eingeführt. Es werden wichtige Anwendungsszenarien und Anforderungen an Positionierungssysteme beschrieben und Indoor- und Outdoor-Systeme voneinander abgegrenzt.

Die Drahtlostechnologie Ultra-wideband eignet sich besonders zur Positionsbestimmung in Innenräumen. Bezugnehmend darauf wird neben regulatorischen Aspekten auch ein internationaler Standard für die drahtlose Entfernungsmessung vorgestellt.

### 2.1 Positionierungssysteme

Positionierungssysteme kommen zum Einsatz, wenn die Aufenthaltsorte von Personen und anderen Objekten ermittelt werden müssen. Heutzutage ist eine Positionsbestimmung in vielen Situationen und Anwendungen erforderlich. Die Tabelle 2.1 führt eine Reihe möglicher Lokalisierungsanwendungen auf und bewertet die nötigen Genauigkeitsanforderungen. Aus diesen Einschätzungen nach Ingram et al. [13] geht hervor, dass für die meisten Anwendungen eine Genauigkeit von 0,3 m bei der Positionsbestimmung ausreichend ist. [3, 13]

Aufgrund erheblicher Unterschiede in den Anforderungen und Eigenschaften der Umgebungen werden Lokalisationssysteme in Outdoor- und Indoor-Positionierungssysteme unterteilt. Nachfolgend werden beide Anwendungsdomänen charakterisiert.

#### 2.1.1 Outdoor-Positionierungssysteme

Zur Positionsbestimmung sind allgemein Global Navigation Satellite Systeme (GNSS) bekannt und stark verbreitet. Dabei senden Satelliten Positionssignale, woraus ein spezieller

Tabelle 2.1: Genauigkeitsanforderungen an verschiedene Lokalisierungsanwendungen

Anwendung	Genauigkeit
Automatisierte Materialflussanlagen	0,5 cm
Wegweisung für Blinde	1 cm
Gebäudevermessung	1 cm
Werkzeugpositionierung	1 cm
Navigation von Robotern innerhalb von Gebäuden	8 cm
Formationsflug von Fluggeräten	10 cm
Freizeit und Spielzeug	10 cm
Häuser- und Straßenschluchten (Off-Road)	30 cm
Häuser- und Straßenschluchten (maritim)	50 cm
Häuser- und Straßenschluchten (sonstiges)	80 cm
Personenführungsanlagen in Ausstellungen	1 m
Ortung von Waren und Gegenständen	1 m
Gefahrenwarnungen	1 m
Navigation von Fußgängern	1 m
Ortung von Arbeitern innerhalb von Gebäuden	1 m
Häuser- und Straßenschluchten (Schienenverkehr)	1 m
Präzises Landen von Fluggeräten	1 m
Zugangskontrolle	3 m
Standortbasierte Dienste	3 m
Präsentation von ortsbezogenen Informationen	10 m
Zug-/Flug-/Businformationen	30 m
Werbung	100 m

Quelle: In Anlehnung an [13]

Empfänger seine eigene Position berechnen kann. Das bekannteste Satellitennavigationssystem ist das Global Positioning System (GPS). Weitere globale Satellitensysteme sind GLONASS, Galileo und Beidou. [16]

Viele Umweltfaktoren nehmen Einfluss auf die Genauigkeit von GPS-Systemen. Häufig wird in diesem Zusammenhang von Häuser- und Straßenschluchten gesprochen, da Gebäude, dichte Bewaldung etc. Satellitensignale blockieren, wodurch die Genauigkeit von GNSS stark abnimmt. Die bekannten Satellitennavigationssysteme sind daher für Anwendungsfälle in Innenräumen nicht geeignet. Auch in urbanen oder bewaldeten Gebieten kann die Leistungsfähigkeit eingeschränkt sein. [16]

Die auf den Consumer-Markt verbreiteten GPS-Systeme bieten in der Regel eine Genauigkeit von 15 m - 20 m. Diese Auflösung ist für bestimmte Anwendungsszenarien nicht immer ausreichend. Durch zusätzliche hardwaregestützte Systeme kann die Genauigkeit von GPS-Systemen weiter gesteigert werden. Verschiedene GPS-Typen wie D-GPS bieten Genauigkeiten zwischen 50 cm - 10 m. Hochpreisige Echtzeitkinematik-Systeme können Genauigkeiten von 2 cm - 3 cm erreichen. Diese GPS-Typen setzen zusätzliche Bodenstationen voraus, von denen spezielle Korrektursignale gesendet werden. Dies ist jedoch kostenintensiv und aufwendig. [27]

### 2.1.2 Indoor-Positionierungssysteme

Als Indoor Positioning System (IPS) wird ein System bezeichnet, das kontinuierlich und in Echtzeit die Position von Personen oder Objekten in einer Innenraumumgebung bestimmt. In Privathaushalten können IPS zu verschiedenen Zwecken eingesetzt werden, z. B. zum Verfolgen und Finden von Gegenständen oder zur Ortungs-Unterstützung von pflegebedürftigen Menschen. In öffentlichen Gebäuden kann ein IPS für verschiedene nützliche Anwendungen eingesetzt werden, z. B. zur Bereitstellung von Indoor-Navigationssystemen für blinde und sehbehinderte Menschen, zur Unterstützung von Touristen oder zur Verfolgung von Kindern in überfüllten Einrichtungen. Darüber hinaus können mithilfe des IPS Rettungseinsätze der Polizei und Feuerwehr besser koordiniert werden. Beispielsweise ist die Ortung von Feuerwehrleuten in einem brennenden Gebäude wichtig für die Einsatzleitung und für die Planung weiterer Rettungsmaßnahmen. Ein weiterer Anwendungsfall ergibt sich in der Überwachung und Verfolgung von Gegenständen und Geräten sowie der Verhinderung von Diebstählen. In der Industrie und Logistik nehmen IPS eine wichtige Rolle ein. Mit der fortschreitenden Entwicklung von Automatisierung und Steuerung werden Industrieroboter und andere automatisierte Materialflussanlagen in intelligenten Fabriken durch IPS navigiert. [3]

Die Ortung in Innenräumen unterscheidet sich in vielen Merkmalen von der Ortung im Freien. Durch vergleichsweise viele Objekte (z. B. Wände, Personen, Gegenstände), die Signale reflektieren oder dämpfen, ist eine drahtlose Positionsbestimmung in Innenräumen oftmals komplexer. Durch diese Objekte treten in Innenräumen typischerweise Funkverbindungen ohne direkten Sichtkontakt zwischen Sender und Empfänger (engl. non-line-of-sight (NLoS)) auf, bei denen sich die Signale nicht auf direktem Weg von einem Sender zu einem Empfänger bewegen können, was zu inkonsistenten Zeit- und Signalverzögerungen auf der Empfängerseite führt. Weiterhin leidet die Signalqualität in

Innenräumen, da aufgrund vieler Störquellen in der Umgebung wie z. B. Mobilfunkgeräte, Bluetooth-Geräte, WiFi-Geräte aber auch Mikrowellenherde und Leuchtstoffröhren die Signalstärke leicht schwanken kann. [3]

Typischerweise erfordern Positionierungsanwendungen im Innenbereich eine höhere Präzision und Genauigkeit als eine Positionsbestimmungen im Freien, um die relativ kleinen Bereiche und die vorhandenen Hindernisse in Innenräumen ausreichend gut abbilden zu können. Wiederrum vorteilhaft können die bekannten Strukturen und geringere Dynamik von Innenräumen sein, die die Positionsbestimmung innerhalb eines kleinen Bereiches vereinfachen. [3]

## 2.2 Ultra-wideband

Ein UWB-Signal wird im Allgemeinen als ein Signal mit einer Teilbandbreite von mehr als 20 % oder einer absoluten Bandbreite von mindestens 500 MHz definiert. Eines der Hauptmerkmale von UWB-Signalen beruht darauf, dass sie ein deutlich breiteres Frequenzband als herkömmliche Signale belegen. Daher müssen sich UWB-Anwendungen das vorhandene Frequenzspektrum mit den anderen etablierten Signalen teilen. Folglich gelten für UWB-Systeme bestimmte Vorschriften. [20]

In den nachfolgenden Abschnitten wird daher neben der genaueren Beschreibung der Technologie auch auf Regulierungsvorschriften für UWB-Systeme eingegangen.

### 2.2.1 Technologie

Ein Signal wird als UWB-Signal definiert, wenn es eine absolute Bandbreite von mindestens 500 MHz oder eine Teilbandbreite – auch als relative Bandbreite bezeichnet – von mehr als 20 % besitzt. Die absolute Bandbreite  $B$  wird als Differenz zwischen der oberen Frequenz  $f_H$  und der unteren Frequenz  $f_L$  berechnet. [20]

$$B = f_H - f_L \tag{2.1}$$

An der Mittelfrequenz  $f_c$  weist das Signal seine maximale Leistungsdichte auf. Die untere und obere Frequenz wird an den Orten bestimmt, an denen die spektrale Leistungsdichte

10 dB unterhalb von  $f_c$  liegt. Dieser Zusammenhang ist in Abbildung 2.1 schematisch dargestellt. [26, 20]

Die relative Bandbreite  $B_{frac}$  ist hingegen definiert als

$$B_{frac} = \frac{B}{f_c}, \quad (2.2)$$

wobei die Mittelfrequenz  $f_c$  folgendermaßen angegeben wird:

$$f_c = \frac{f_H + f_L}{2} \quad (2.3)$$

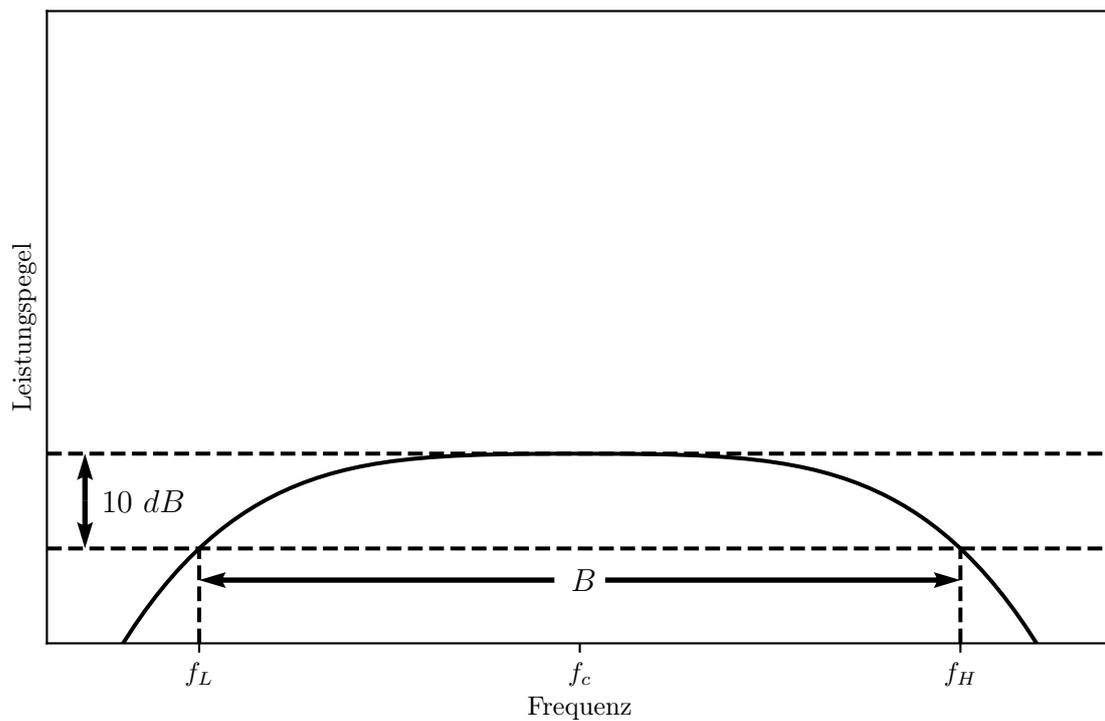


Abbildung 2.1: Untere, mittlere und obere Frequenz eines UWB-Signals

Quelle: In Anlehnung an [26]

Aufgrund der großen Bandbreite zeichnen sich UWB-Systeme durch sehr kurze Wellenformen aus, die normalerweise in der Größenordnung einer Nanosekunde liegen. Üblicherweise sendet ein UWB-System ultrakurze Impulse mit einem niedrigen Tastverhältnis. Das Verhältnis zwischen dem Zeitpunkt der Impulsübertragung und der durchschnittlichen Zeit zwischen zwei aufeinanderfolgenden Übertragungen wird somit klein gehalten. [20]

Die große Bandbreite sorgt weiterhin für eine Reihe von Vorteilen für Ortungs- und Kommunikationsanwendungen. [20, 26]

Die wichtigsten Eigenschaften lassen sich wie folgt zusammenfassen:

- Infolge des breiten Frequenzspektrums, das auch niedrige Frequenzen umfasst, können die Signale leichter Hindernisse durchdringen.
- Durch die kurzen Sendeimpulse ist die zeitliche Auflösung von UWB-Signalen sehr hoch, wodurch sich sehr genaue Entfernung- und Positionsmessungen realisieren lassen.
- Die Datenkommunikation ist mit hohen Datenraten möglich.
- Die Signalform ähnelt dem Grundrauschen. Die Sendeleistung wird auf den gesamten Frequenzbereich verteilt. Somit ist nur schwer feststellbar, dass eine UWB-Kommunikation stattfindet, was die Abhörsicherheit erhöht.
- Die Energieaufnahme und die Kosten eines UWB-Transceivers sind gering.

Da die UWB-Signale einen sehr großen Teil des Frequenzspektrums beanspruchen, kommt es zu Überlagerungen mit etablierten schmalbandigen Drahtlossystemen. In diesen Systemen dürfen keine Störungen durch UWB-Signale verursacht werden. Ein Beispiel für die Frequenzverteilung verschiedener Drahtlossysteme ist in Abbildung 2.2 dargestellt. Häufig ist für UWB-Kommunikation ein Grenzwert von  $-41,3 \text{ dBm MHz}^{-1}$  vorgeschrieben, der ebenfalls in der Abbildung eingezeichnet ist. Eine ähnliche Leistungsbeschränkung besteht auch für andere unbeabsichtigte Strahler wie Fernsehgeräte oder Computermonitore. [20]

### 2.2.2 Regulation

Wenn UWB-Systeme ohne Einschränkungen im Frequenzbereich anderer Systeme senden dürften, könnten diese Systeme durch die UWB-Signale gestört werden. Ein UWB-Sender

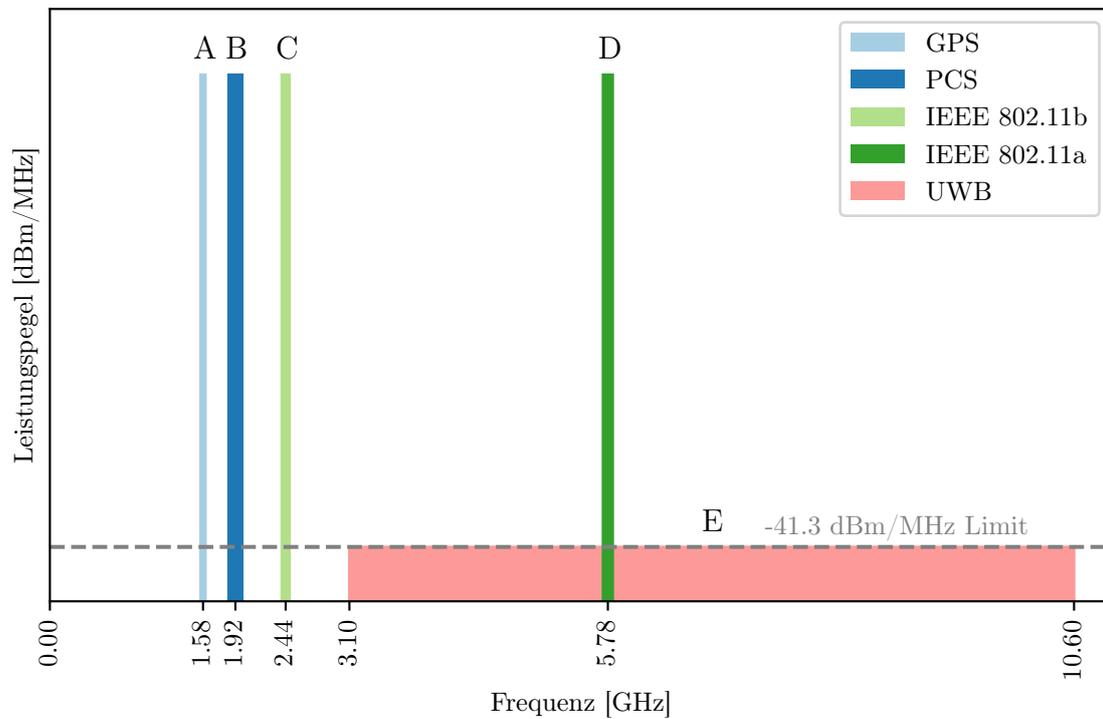


Abbildung 2.2: Frequenzverteilung verschiedener Drahtlosysteme.

A: Global Positioning System (GPS) (1,56-1,61 GHz); B: Personal Communication System (PCS) (1,85-1,99 GHz); C: Mikrowellenherde, schnurlose Telefone, Bluetooth, Wi-Fi IEEE 802.11b (2,4-2,48 GHz); D: Wi-Fi IEEE 802.11a (5,725-5,825 GHz); E: Ultra-wideband (UWB) (3,1-10,6 GHz).

Die Leistungspegel der verschiedenen Systeme sind nicht maßstabsgetreu dargestellt.

Quelle: In Anlehnung an [20]

muss daher bestimmte Anforderungen erfüllen, um die Funktionalität der anderen Systeme nicht zu beeinflussen. [20]

In den USA hat die Federal Communications Commission (FCC) als Regulierungsbehörde im Jahr 1998 eine Spezifikation für Ultra-wideband eingeführt. Seit dem Jahr 2002 ist die eingeschränkte Nutzung von UWB-Systemen erlaubt. Nach dieser Verordnung darf die spektrale Leistungsdichte den Grenzwert von  $-41,3 \text{ dBm MHz}^{-1}$  für Frequenzbereiche von 3,1 GHz - 10,6 GHz nicht überschreiten und muss außerhalb dieses Bandes je nach Anwendung noch niedriger sein. Daraus ergibt sich ein für die meisten UWB-Systeme nutzbares Spektrum von 7,5 GHz. [20]

Trotz dieser Einschränkung liegt ein wichtiger Vorteil darin, dass die Nutzung von Ultra-wideband lizenzfrei ist und davon profitiert, dass die UWB-Kommunikation von der Allgemeinheit ohne Genehmigung in einem bestimmten Leistungsbereich mit einer Reihe von Frequenzen genutzt werden kann. Nachdem die FCC die Verwendung von UWB-Signalen in den USA legalisiert hat, wurden erhebliche Fortschritte bei der Entwicklung und der Standardisierung von UWB-Systemen erzielt. [26, 20]

In Deutschland ist der Frequenzbereich 30 MHz - 10,6 GHz erstmals seit dem 16. Januar 2008 von der Bundesnetzagentur für jedermann freigegeben. Die Bundesnetzagentur ist als deutsche Regulierungsbehörde unter anderem zuständig für die Zuteilung verschiedener Frequenzbänder. [7]

In der Allgemein-zuteilung schränkt die Bundesnetzagentur die Verwendung des UWB-Frequenzbereich präziser ein. Zwar ist der Frequenzbereich grundsätzlich für jedermann kostenlos und genehmigungsfrei nutzbar, allerdings müssen die Auflagen der Allgemein-zuteilung berücksichtigt werden. Je nach Anwendungsgebiet und Einsatzort werden Beschränkungen bei der maximalen Sendeleistung und/oder des maximalen Sendezeitanteils gemacht. Die höchsten Sendeleistungen von  $-41,3 \text{ dBm MHz}^{-1}$  sind im Frequenzbereich von 6,0 GHz - 8,5 GHz erlaubt. [8]

### 2.3 Einsatzgebiete

Die Ultra-wideband Technologie eignet sich besonders zur Echtzeit-Positionsbestimmung. Neben dem klassischen Anwendungsfall sind auch weitere spezifischere Szenarien wie z. B. Geofencing, Spatial Awareness etc. denkbar. Darüber hinaus kann Ultra-wideband auch allgemein zur drahtlosen Datenübertragung eingesetzt werden.

UWB-Lokalisationssysteme sind nicht zwangsweise auf Innenräume beschränkt. Im Rahmen von zukunftsweisenden Projekten zur Quartiersmobilität ist es denkbar, dass einzelne Stadtquartiere infrastrukturell mit der Ultra-wideband Technologie ausgestattet werden. Im Außenbereich sind auf Ultra-wideband basierende Lokalisationssysteme insbesondere sinnvoll, wenn aufgrund einer dichten städtischen Bebauung keine präzise GPS-Navigation möglich ist.

Neben UWB-Lokalisationssystemen sind zudem Indoor-Positionierungssysteme bekannt, die auf Radio Frequency Identification (RFID), WiFi, Bluetooth, ZigBee oder anderen optischen oder akustischen Verfahren basieren. Diese Methoden sind in der Regel jedoch wesentlich ungenauer und für präzise Ortungen nicht geeignet. Nachteilig im Einsatz von Ultra-wideband ist, dass nicht auf eine bestehende Infrastruktur – wie es z. B. bei WiFi der Fall ist – zurückgegriffen werden kann. Die Anwendung von UWB-Lokalisationssystemen ist also nur punktuell möglich. [3]

### 2.4 IEEE 802.15.4

Der internationale Standard IEEE 802.15.4 spezifiziert die Bitübertragungsschicht (Physical Layer) und Sicherungsschicht (Data Link Layer) für die drahtlose Entfernungsmessung. Der Standard definiert zwei Formate für Kommunikationssignale: Impulse Radio Ultra-wideband (IR-UWB) und Chirp Spread Spectrum (CSS) Signale. Während das CSS-Format für die Datenkommunikation geeignet ist, bietet IR-UWB die Möglichkeit zur präzisen Entfernungsmessung. [26]

Im Weiteren wird daher das IR-UWB Signalformat genauer betrachtet.

#### 2.4.1 IEEE 802.15.4 Paketaufbau

Die Abbildung 2.3 stellt den Aufbau eines Nachrichtenpakets nach dem IEEE 802.15.4 Standard dar.

##### Preamble

Eine Preamble wird in Netzwerkprotokollen verwendet, um den Empfänger zu synchronisieren und über die Ankunft eines Nachrichtenpakets zu informieren. Gemäß des IEEE

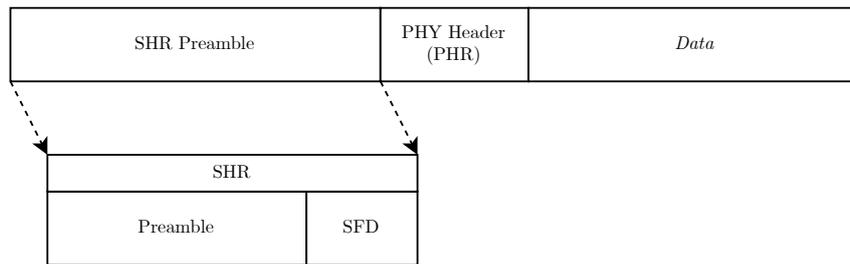


Abbildung 2.3: Darstellung der IEEE 802.15.4 Paketstruktur

Quelle: In Anlehnung an [20]

802.15.4 Standards kann eine Preamble-Länge von 16, 64, 1024 oder 4096 Symbolen gewählt werden. [26]

Die Länge der Preamble wird an die Anforderungen der Anwendung angepasst. Eine lange Preamble erhöht die Zuverlässigkeit bei weiten Kommunikationsdistanzen oder schwachen Signalpegeln. Eine kürzere Länge bietet hingegen Vorteile durch einen geringeren Energieverbrauch und eine kürzere Nachrichtenlänge. Dadurch wird die Kanalbelegung reduziert, wodurch mehr Teilnehmer den Kommunikationskanal nutzen können. Der Standard schlägt für Positionierungssysteme allgemein eine Preamble-Länge von 1024 Symbolen vor. [26, 1]

### Start Frame Delimiter (SFD)

Mit dem Start Frame Delimiter (SFD) wird das Ende der Preamble und der Beginn des Nachrichtenheaders signalisiert. Der SFD ist eine kurze Sequenz aus 8 oder 64 Symbolen. Zur präzisen Entfernungsmessung ist es entscheidend, dass die Ankunftszeit einer Nachricht und die Antwortzeit exakt gemessen wird. Der SFD stellt dazu ein schmales Trigger dar, der als Signal zum Starten bzw. Stoppen der Zeitmessung verwendet wird.

### PHY Header (PHR)

Der Header besteht aus Feldern, die die Datenrate, die Framelänge, das Ranging-Bit, die Preamble-Länge sowie die Bits zur Fehlerkorrektur und -erkennung angeben. Die Länge

beträgt 19 Bytes. Ein gesetztes Ranging-Bit zeigt an, dass es sich um eine Nachricht zur Entfernungsmessung handelt.

### **Data**

Das Datenfeld enthält die eigentlichen Nutzdaten der Nachricht. Die im UWB-Lokalisationssystem eingesetzten Nachrichten sind in Kapitel 4 dokumentiert.

## **2.5 Terminologie**

In dieser Arbeit werden die nachfolgenden Bezeichnungen für die verschiedenen Teilnehmer eines UWB-Lokalisationssystems eingeführt. Im Allgemeinen werden diese Teilnehmer auch als UWB-*Beacon* bezeichnet. Der Begriff UWB-*Anchor* (dt. Anker) spezifiziert genauer, dass es sich um einen ortsfesten Teilnehmer mit bekannter Position handelt, der als Referenzpunkt im Lokalisationssystem verwendet wird. Der Begriff UWB-*Tag* (dt. Anhänger) kennzeichnet wiederum das bewegliche zu lokalisierende Ziel mit unbekannter Position.

### 3 Positionsbestimmung

Zur Bestimmung der Position eines Knotens in einem drahtlosen Netzwerk werden Signale zwischen einer Reihe von (Referenz-)Knoten ausgetauscht. Um aus diesen Signalen die Position ermitteln zu können, existieren zwei grundlegende Ansätze.

Die Positionsbestimmung kann direkt aus den zwischen den Knoten gesendeten Signalen erfolgen, was als direkter Ansatz bezeichnet wird.

Weiterhin eignet sich ein zweistufiger Ansatz, bei dem zuvor bestimmte Parameter aus dem Signal extrahiert werden, die dann zur Positionsbestimmung weiterverwendet werden (vgl. Abbildung 3.1). [26]

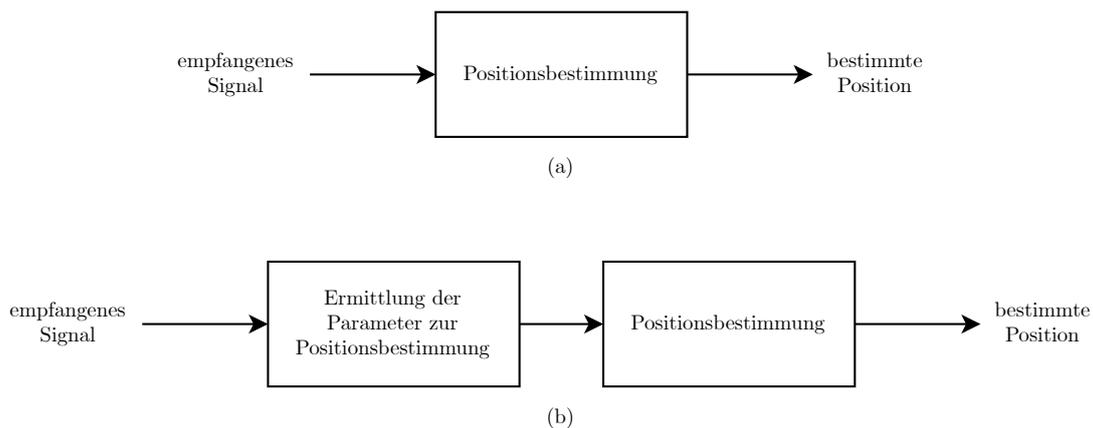


Abbildung 3.1: Direkter Ansatz (a) und zweistufiger Ansatz (b) zur Positionsbestimmung

Quelle: In Anlehnung an [20]

Da bei Verwendung des zweistufigen Ansatzes die Parameter irrtümlich aus unerwünschten Rauschsignalen extrahiert werden können, ist ein ausreichend hohes Signal-Rausch-Verhältnis (SNR) oder eine hohe Signalbandbreite erforderlich.

Der zweistufige Ansatz ist verglichen mit dem direkten Ansatz weniger komplex. In der Praxis verwenden daher viele Systeme den zweistufigen Ansatz. [26, 20]

Im Folgenden werden in diesem Kapitel verschiedene Signalparameter vorgestellt, die sich zu Positionsbestimmung eignen. Zudem werden gängige Verfahren zur Positionsbestimmung beschrieben.

## 3.1 Signalparameter zur Positionsbestimmung

Aus den Signalparametern werden zunächst einfache Eigenschaften abgeleitet, aus denen anschließend eine Position bestimmt werden kann. Die Parameter liefern z. B. Entfernungen oder Winkel.

### 3.1.1 Received Signal Strength Indicator (RSSI)

Der Received Signal Strength Indicator (RSSI) ist eine Maßeinheit für den Leistungspegel eines empfangenen Funksignals und bietet damit einen einfachen Anhaltspunkt zur Messung von Entfernungen.

Der Grundgedanke eines RSSI-basierten Ansatzes ist der bekannte Zusammenhang zwischen der Verlustleistung und der Entfernung. Der Leistungsverlust eines elektromagnetischen Signals zwischen Sender und Empfänger wird mit dem Pfadverlust (engl. path loss) beschrieben. Der Pfadverlust ist dabei direkt proportional zur Distanz zwischen Sender und Empfänger. Eine größere Entfernung zwischen den zwei Knoten führt zu einer höheren Verlustleistung, also zu einem höheren Pfadverlust. [20]

Vorausgesetzt, dass die Sendeleistung des sendenden Knotens bekannt ist, kann ein empfangender Knoten die eingehende Sendeleistung messen und daraus die Distanz zwischen Sender und Empfänger ableiten. [20]

Aufgrund von Ungenauigkeiten sowohl bei den Messungen der Signalstärke als auch in der Bestimmung des Verhältnisses zwischen Entfernung und Pfadverlust sind die Entfernungsschätzungen mit größeren Fehlern behaftet. Es gibt zwei wesentliche Phänomene, die den Pfadverlust eines Signales beeinflussen.

Ein Effekt ist die Mehrwegeausbreitung (engl. multipath propagation), bei dem die Signalanteile eines Funksignals auf mehreren verschiedenen Wegen vom Sender zum Empfänger gelangen, wodurch unterschiedliche Werte für den Pfadverlust gemessen werden. Das zweite Phänomen ist die Abschattung (engl. shadowing). Ursächlich dafür ist eine sich ändernde Umgebung bei der Ausbreitung eines Signals über größere Entfernungen.

Durch diese Veränderung der Empfangsbedingungen werden Ungenauigkeiten bei der Messung des Pfadverlusts induziert.

Zwar existieren für beide Phänomene analytische und stochastische Methoden, um die damit verbundenen Probleme zu umgehen. Dennoch ist in UWB-Drahtlossystemen bereits bei geringen Distanzen eine hohe Standardabweichungen in der RSSI-basierten Entfernungsabschätzung vorhanden. [26]

Insgesamt können RSS-Messungen keine sehr genauen Entfernungsmessung für UWB-Systeme liefern.

#### 3.1.2 Time of Arrival (ToA)

Mit dem Time of Arrival (ToA) Verfahren wird die Entfernung zwischen zwei Knoten bestimmt, indem die Signallaufzeit zwischen Sender und Empfänger gemessen wird.

Zum Abgleich der Signallaufzeiten erfordert das Verfahren eine synchronisierte Uhr über die beteiligten Knoten hinweg oder den Austausch der Zeitstempel des Nachrichteneingangs bzw. des Nachrichtenversands. Aus diesen Zeitinformationen kann die Distanz zwischen den zwei Knoten ermittelt werden. [20]

Die Distanz  $s$  zwischen zwei Knoten ist dabei das Produkt aus der konstanten Lichtgeschwindigkeit  $c$  und der gemessenen Signallaufzeit  $t_{prop}$ .

$$\underbrace{s}_{\text{Distanz}} = \underbrace{c}_{\text{Lichtgeschwindigkeit}} \cdot \underbrace{t_{prop}}_{\text{Signallaufzeit}} \quad (3.1)$$

Auf Grundlage der Distanzen zwischen dem beweglichen Tag und den bekannten Positionen der Referenzknoten kann die genaue Position des Tags errechnet werden. Die gesuchte Position befindet sich dabei auf Kreisen (bzw. Kugelschalen) mit den Radien der jeweiligen Entfernungen. Diese Möglichkeit der Positionsbestimmung ist im nachfolgenden Abschnitt 3.2.2 Trilateration genauer erläutert.

Für den Austausch der nötigen Zeitstempelinformationen stehen zwei gängige Protokolle zur Verfügung:

### Single-Sided Two-Way Ranging (SS-TWR)

Mit dem Single-Sided Two-Way Ranging (SS-TWR) wird die Paketumlaufzeit (engl. round-trip time (RTT)) einer Nachricht von einem sendenden Knoten zu einem empfangenden Knoten gemessen. Das Messergebnis wird dem ursprünglichen Sender zurückgeschickt. [11]

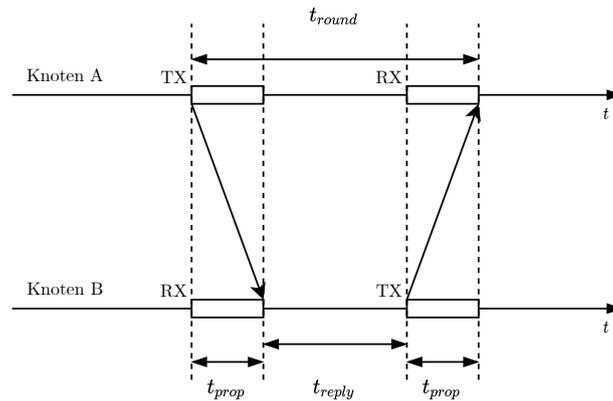


Abbildung 3.2: Single-Sided Two-Way Ranging (SS-TWR)

Quelle: In Anlehnung an [11]

Die Funktionsweise von SS-TWR wird in Abbildung 3.2 dargestellt. Knoten A leitet den Nachrichtenaustausch ein, woraufhin Knoten B antwortet, um die Messung abzuschließen. Jeder Knoten bestimmt die genauen Sende- bzw. Empfangszeitstempel, sodass die Zeiten  $t_{round}$  und  $t_{reply}$  durch die einfache Subtraktion dieser Zeitstempel bekannt sind. Die ungefähre Signallaufzeit  $\hat{t}_{prop}$  kann anschließend mit der folgenden Gleichung berechnet werden:

$$\hat{t}_{prop} = \frac{1}{2}(t_{round} - t_{reply}) \quad (3.2)$$

Die Zeiten  $t_{round}$  und  $t_{reply}$  werden voneinander unabhängig durch die Knoten A und B mithilfe ihrer internen Oszillatoren gemessen. Beide Oszillatoren haben verschiedene Abweichungen  $e_A$  und  $e_B$  von ihrer Nennfrequenz, sodass mit größer werdendem  $t_{reply}$  der Messfehler ebenfalls ansteigt.

Dieser Fehler in der gemessenen Signallaufzeit setzt sich wie folgt zusammen:

$$error = \hat{t}_{prop} - t_{prop} \approx \frac{1}{2}(e_B - e_A) \cdot t_{reply} \quad (3.3)$$

In Tabelle 3.1 sind einige typische Werte für die Größen dieses Messfehlers aufgeführt. Es ist zu beachten, dass die Antwortzeit  $t_{reply}$  nicht nur die reine Reaktionszeit (RX-zu-TX-Durchlaufzeit) der UWB-Hardware ist, sondern auch die Nachrichtenlänge umfasst. [26]

Tabelle 3.1: Typische zeitbedingte Fehler bei der SS-TWR Signallaufzeitmessung

$t_{reply}$ \ Taktversatz	2 ppm	5 ppm	10 ppm	20 ppm	40 ppm
100 $\mu$ s	0,1 ns	0,25 ns	0,5 ns	1 ns	2 ns
200 $\mu$ s	0,2 ns	0,5 ns	1 ns	2 ns	4 ns
500 $\mu$ s	0,5 ns	1,25 ns	2,5 ns	5 ns	10 ns
1 ms	1 ns	2,5 ns	5 ns	10 ns	20 ns
2 ms	2 ns	5 ns	10 ns	20 ns	40 ns
5 ms	5 ns	12,5 ns	25 ns	50 ns	100 ns

Quelle: In Anlehnung an [11, 26]

Der Tabelle 3.1 ist zu entnehmen, dass mit zunehmendem  $t_{reply}$  und zunehmendem Taktversatz der Fehler in der Signallaufzeitbestimmung ansteigt. Ab einem gewissen Punkt wird der Fehler so groß, dass die Bestimmung sehr ungenau wird. Unter der vereinfachten Annahme, bei der sich ein Funksignal mit einer Geschwindigkeit von  $300\,000\text{ km s}^{-1}$  ausbreitet, entspricht ein Fehler von 1 ns einer Abweichung von 30 cm in der bestimmten Entfernung. [11]

Aus diesem Grund wird das SS-TWR in Praxisanwendungen üblicherweise nicht eingesetzt. Dennoch bietet das Verfahren durch seine Einfachheit Vorteile, wenn sehr genaue Oszillatoren verwendet werden und das System auf kurze Kommunikationsdistanzen beschränkt ist. [11]

### Double-Sided Two-Way Ranging (DS-TWR)

Das Double-Sided Two-Way Ranging (DS-TWR) stellt eine Erweiterung des einfachen SS-TWR dar. Durch die Ausführung von zwei Messungen der Paketumlaufzeit und der Kombination der gemessenen Zeiten können Ergebnisse erzielt werden, die auch bei langen Antwortzeiten einen geringen Fehler aufweisen. [11]

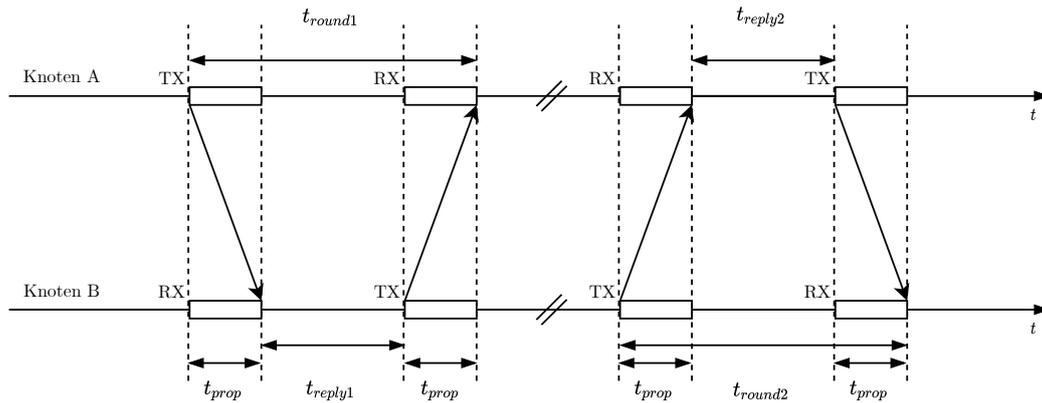


Abbildung 3.3: Double-Sided Two-Way Ranging (DS-TWR) mit vier Nachrichten

Quelle: In Anlehnung an [11]

Die Funktionsweise von DS-TWR wird in Abbildung 3.3 dargestellt. Der Knoten A leitet die erste Paketumlaufzeitmessung ein, woraufhin Knoten B antwortet, um die Messung abzuschließen und daraufhin die zweite Paketumlaufzeitmessung startet. Mit der abschließenden Antwort von Knoten A wird die zweite Messung beendet.

Jeder Knoten bestimmt die genauen Sende- bzw. Empfangszeitstempel, sodass die Zeiten  $t_{round1}$  und  $t_{reply1}$  für den ersten Messdurchgang und die Zeiten  $t_{round2}$  und  $t_{reply2}$  des zweiten Messdurchgangs bekannt sind. Durch dieses Verfahren werden beide Zeiten gleichermaßen einmal von Knoten A und einmal von Knoten B bestimmt. Die sich daraus ergebende ungefähre Signallaufzeit kann folgendermaßen berechnet werden:

$$\hat{t}_{prop} = \frac{t_{round1} \cdot t_{round2} - t_{reply1} \cdot t_{reply2}}{t_{round1} + t_{round2} + t_{reply1} + t_{reply2}} \quad (3.4)$$

Das Protokoll kann im Nachrichtenbedarf dahingehend optimiert werden, dass anstatt von vier Nachrichten nur drei Nachrichten notwendig sind, um einen Austausch mit zwei Paketumlaufzeitmessungen durchzuführen. Der optimierte Ablauf ist in Abbildung 3.4 dargestellt. Die Antwort des Knoten B zum Abschluss des ersten Messdurchgangs kann mit der Initiierung des zweiten Messdurchgangs vereint werden. [11]  
 Ein geringer Nachrichtenbedarf optimiert das System hinsichtlich des Zeit- und Energiebedarfs.

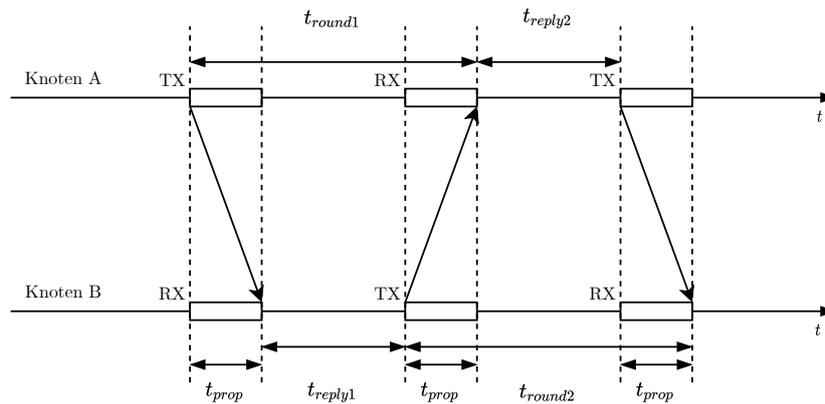


Abbildung 3.4: Double-Sided Two-Way Ranging (DS-TWR) mit drei Nachrichten

Quelle: In Anlehnung an [11]

Durch die Messung der Antwortzeiten  $t_{reply1}$  und  $t_{reply2}$  sowohl auf Knoten A als auch auf Knoten B, kann der Zeitversatz beider Oszillatoren kompensiert werden.

In Tabelle 3.2 sind einige typische Fehlergrößen für das DS-TWR Verfahren aufgeführt. Der Fehler wird folgendermaßen berechnet:

$$error = \hat{t}_{prop} - t_{prop} \approx \frac{1}{2}(e_B - e_A) \cdot \frac{1}{2}(t_{reply2} - t_{reply1}) \quad (3.5)$$

Auch bei Taktversätzen von 20 ppm liegt der induzierte Fehler im Bereich weniger Piko-sekunden. Bei dieser Fehlergröße ist die Genauigkeit der Bestimmung der Nachrichtenan-kunftszeit an den einzelnen Knoten ein weitaus wichtigerer Faktor für den Gesamtfehler, als der durch die Oszillatoren verursachte Fehler. [11]

Tabelle 3.2: Typische zeitbedingte Fehler bei der DS-TWR Signallaufzeitmessung

Taktversatz $t_{reply2} - t_{reply1}$	2 ppm	5 ppm	10 ppm	20 ppm	40 ppm
1 $\mu$ s	0,0005 ns	0,001 25 ns	0,0025 ns	0,005 ns	0,01 ns
2 $\mu$ s	0,001 ns	0,0025 ns	0,005 ns	0,01 ns	0,02 ns
5 $\mu$ s	0,0025 ns	0,006 25 ns	0,0125 ns	0,025 ns	0,05 ns
10 $\mu$ s	0,005 ns	0,0125 ns	0,025 ns	0,05 ns	0,1 ns
100 $\mu$ s	0,05 ns	0,125 ns	0,25 ns	0,5 ns	1 ns
200 $\mu$ s	0,1 ns	0,25 ns	0,5 ns	1 ns	2 ns

Quelle: In Anlehnung an [11, 26]

In der Praxis wird DS-TWR daher häufig in ToA Anwendungen eingesetzt, auch wenn gegenüber SS-TWR die Komplexität im Zeit- und Rechenbedarf höher ist. Der Vorteil der Kompensation des Zeitversatzes gibt eine hohe Flexibilität bei der Gestaltung von Anwendungsfällen und Anwendungsszenarien. [11]

### 3.1.3 Time Difference of Arrival (TDoA)

Als Time Difference of Arrival (TDoA) wird die Differenz in der Ankunftszeit zweier Signale, die sich zwischen dem Sender und zwei Referenzknoten ausbreiten, bezeichnet. Durch die Multiplikation der Zeitdifferenz mit der Lichtgeschwindigkeit kann eine Hyperbel (bzw. ein Hyperboloid) bestimmt werden, auf der sich die Position des Senders befindet. [26]

Der Vorteil von TDoA gegenüber ToA liegt darin, dass keine zeitliche Synchronisation zwischen den Anchors und dem Tag erforderlich ist oder keine Zeitstempel ausgetauscht werden müssen. Die Anchors müssen im TDoA Verfahren jedoch synchronisiert sein, damit die Zeitunterschiede in den Ankunftszeiten exakt bestimmt werden können. [26]

Im Abschnitt 3.2.3 zur Hyperbelnavigation erfolgt eine genauere Beschreibung, wie die Position des Tags aus den Zeitdifferenzen errechnet werden kann.

### 3.1.4 Angle of Arrival (AoA)

Im Gegensatz zu den zuvor vorgestellten Signalparametern, aus denen sich Informationen über die Entfernung zweier Knoten ableiten lassen, liefert eine Angle of Arrival (AoA) Messung Informationen über die Richtung eines ankommenden Signals. Es wird folglich der Winkel zwischen den beiden Knoten bestimmt. [26]

Üblicherweise ist hierzu ein Antennenarray bestehend aus zwei Antennen nötig. Die Winkelinformation wird anhand des Antennenarrays durch die Messung des Unterschiedes in den Ankunftszeiten eines ankommenden Signals an den einzelnen Antennenelementen ermittelt. [26]

Die Berechnung einer Position aus den Winkelinformationen erfolgt mit der Triangulation. Diese Möglichkeit der Positionsbestimmung ist im nachfolgenden Abschnitt 3.2.1 Triangulation genauer erläutert.

## 3.2 Methoden zur Positionsbestimmung aus den Signalparametern

Im Folgenden werden Methoden vorgestellt, mit denen aus den zuvor beschriebenen Signalparametern die Position im zwei- oder dreidimensionalen Raum festgestellt werden kann.

### 3.2.1 Triangulation

Die Triangulation ist ein Verfahren, um aus Winkelinformationen zu bekannten Referenzpunkten, wie sie durch AoA ermittelt werden, eine Position abzuleiten. Für eine Positionsbestimmung im zweidimensionalen Raum werden zwei Referenzknoten benötigt. [26]

Wie in Abbildung 3.5 dargestellt, liefert AoA für jeden Referenzknoten einen Winkel. Zusammen mit diesen Winkeln  $\psi_1$  und  $\psi_2$  und den Positionen  $L_1$  und  $L_2$  der jeweiligen Referenzknoten kann die Position  $P$  des Zielknotens berechnet werden. [26]

Aus der Trigonometrie ergeben sich folgende Gleichungen, aus denen der Ort von Punkt  $P$  hervorgeht:

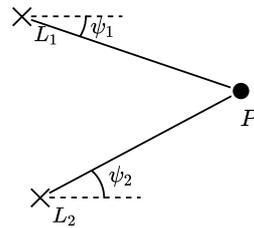


Abbildung 3.5: Geometrische Darstellung der Triangulation

Quelle: In Anlehnung an [20]

$$\tan \psi_1 = \frac{P_y - L_{1,y}}{P_x - L_{1,x}} \quad (3.6)$$

und

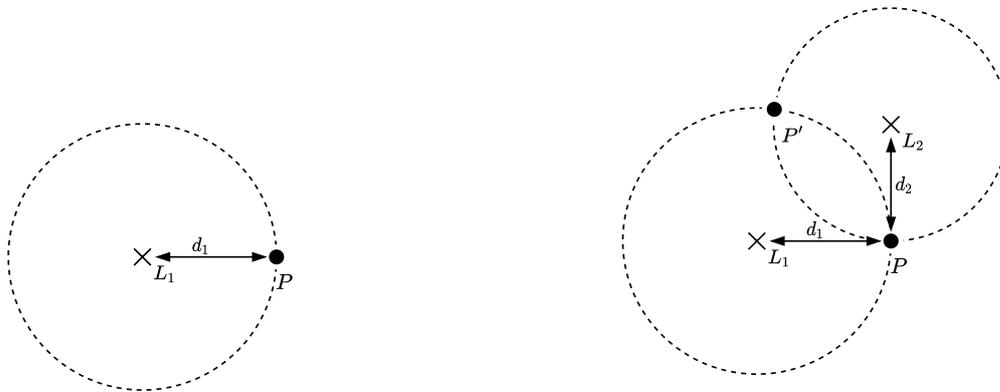
$$\tan \psi_2 = \frac{P_y - L_{2,y}}{P_x - L_{2,x}} \quad (3.7)$$

### 3.2.2 Trilateration

Als Trilateration wird eine Methode bezeichnet, um aus mehreren Entfernungsmessungen, wie sie z. B. RSSI oder ToA liefern, eine Position zu bestimmen. Die Technik basiert darauf, dass sich bei bekannter Entfernung zu einem Punkt der eigene Standort auf einem Kreis (bzw. auf einer Kugelschale) um diesen Punkt befindet. [28]

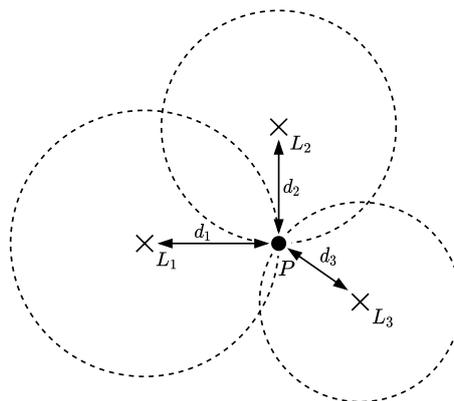
In Abbildung 3.6 wird ein zweidimensionales Trilaterationsproblem geometrisch dargestellt. Es wird die Position  $P$  in Abhängigkeit zu den ortsfesten Referenzknoten  $L_1$ ,  $L_2$  und  $L_3$  gesucht. Die Entfernungen  $d_1$ ,  $d_2$  und  $d_3$  sind bekannt.

Wie aus der Grafik hervorgeht, werden im zweidimensionalen Raum Entfernungsmessungen zu mindestens drei bekannten Punkten benötigt, um die eigene Position eindeutig zu bestimmen. Im dreidimensionalen Raum erhöht sich die Anzahl auf mindestens vier Referenzpunkte. [20]



(a) Mit einer bekannten Entfernung ist jeder Punkt mit dem Abstand  $d_1$  zu  $L_1$  eine potentielle Möglichkeit für Punkt  $P$ .

(b) Mit zwei bekannten Entfernungen wird der Ort für den Punkt  $P$  auf zwei Möglichkeiten eingeschränkt.  $P$  liegt in einem der zwei Schnittpunkte der Kreise um  $L_1$  und  $L_2$ .



(c) Mit drei bekannten Entfernungen ist der Ort von Punkt  $P$  eindeutig.  $P$  liegt im Schnittpunkt der drei Kreise um  $L_1$ ,  $L_2$  und  $L_3$ .

Abbildung 3.6: Geometrische Darstellung der Trilateration

Quelle: In Anlehnung an [28]

Mathematisch kann das Problem als lineares Gleichungssystem formuliert werden. Allgemein betrachtet liegt ein Punkt  $(x, y)$  in der kartesischen Ebene auf einem Kreis mit dem Radius  $r$ , dessen Mittelpunkt  $(M_x, M_y)$  ist, sofern genau eine Lösung für die Gleichung

$$r = \sqrt{(x - M_x)^2 + (y - M_y)^2} \quad (3.8)$$

existiert.

Aus dieser allgemeinen Formulierung können für das zuvor vorgestellte Problem drei Gleichungen für die drei Kreise um die Referenzpositionen abgeleitet werden. Das Gleichungssystem wird gelöst, indem für den Punkt  $P$  eine Lösung gefunden wird, die gleichermaßen für alle drei Gleichungen gilt. Da sich alle drei Kreise in einem Punkt schneiden, hat das Gleichungssystem genau eine Lösung. [28]

$$d_1 = \sqrt{(P_x - L_{1,x})^2 + (P_y - L_{1,y})^2} \quad (3.9a)$$

$$d_2 = \sqrt{(P_x - L_{2,x})^2 + (P_y - L_{2,y})^2} \quad (3.9b)$$

$$d_3 = \sqrt{(P_x - L_{3,x})^2 + (P_y - L_{3,y})^2} \quad (3.9c)$$

In der Praxis ist es nahezu unmöglich, dieses Problem anhand der geometrischen Modellierung zu lösen. Die exakte Lösung des Gleichungssystem würde Positionsangaben und Entfernungsmessungen mit perfekter Genauigkeit voraussetzen. Sobald sich die Kreise nicht in einem einzigen Punkt schneiden, hat das Gleichungssystem keine Lösung.

Auch unter Annahme eines perfekten Systems skaliert dieser mathematische Ansatz zudem nur schlecht über eine größere Anzahl von Referenzknoten. [28]

Die Trilateration kann stattdessen als Optimierungsproblem betrachtet werden. Anstatt direkt den Schnittpunkt der Kreise zu berechnen, wird stattdessen ein Punkt  $X$  gesucht, der die beste Annäherung an die tatsächliche Position  $P$  liefert. Anhand des Punktes  $X$  kann abgeschätzt werden, wie gut dieser Punkt  $P$  ersetzt. Dazu wird die Entfernung zwischen  $X$  zu jedem Anchor  $L_i$  berechnet. Wenn diese Entfernungen genau mit den jeweiligen Entfernungen  $d_i$  übereinstimmen, ist  $X$  identisch zu  $P$ . Je weiter  $X$  von diesen Entfernungen abweicht, desto weiter ist  $X$  von  $P$  entfernt. [28]

Das Ziel des Optimierungsproblems ist einen Punkt  $X$  zu finden, der eine bestimmte Fehlerfunktion minimiert. Jeder Anchor im System stellt eine Fehlerquelle dar, sodass sich zur allgemeinen Berechnung einer Fehlerquelle  $e_i$  folgende Gleichung ergibt:

$$e_i = d_i - \text{dist}(X, L_i) \quad (3.10)$$

Die einzelnen Fehlerquellen können zusammengefasst werden, indem z.B. die mittlere quadratische Abweichung (engl. mean squared error) gebildet wird. Zur Suche eines optimalen  $X$  eignen sich bekannte Algorithmen zur Reduzierung dieser Fehlerfunktion. [28]

### 3.2.3 Hyperbelnavigation

Die Hyperbelnavigation ist ein Verfahren, mit dem sich eine Position mithilfe von TDoA Zeitdifferenzen bestimmen lässt. Im zweidimensionalen Raum werden drei bekannte Referenzknoten benötigt, sodass zwei Zeitdifferenzen ermittelt werden können. [26]

Aus der Geometrie ist bekannt, dass alle Punkte, die zu zwei Orten die gleiche Abstandsdifferenz haben, auf einer Hyperbel liegen. Wie in Abbildung 3.7 dargestellt, resultieren aus zwei TDoA Parametern zwei Hyperbeln, deren Brennpunkte in den Orten der Referenzknoten liegen. Die gesuchte Position liegt im Schnittpunkt der Hyperbeln. [26]

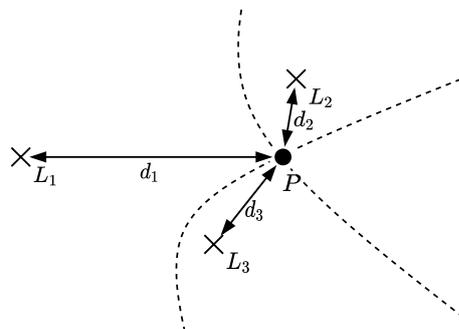


Abbildung 3.7: Geometrische Darstellung der Hyperbelnavigation

Quelle: In Anlehnung an [20]

### 3.3 Dilution of Precision

Die Genauigkeit der Positionsermittlung mithilfe der Trilateration hängt nicht nur von der Sicherheit der Entfernungsmessungen und den bekannten Referenzpositionen ab. Neben diesen Eigenschaften ist auch die geometrische Anordnung der Anchors untereinander und in Abhängigkeit zur Position des zu bestimmenden Ortes von Bedeutung. Mit der Dilution of Precision (DOP) wird die mögliche Streubreite der Positionsmessung angegeben. [16]

Die Streubreite der Positionsbestimmung ist gering, wenn die Standorte der Teilnehmer in rechten Winkeln zueinander liegen. Schneiden sich die Kreise oder Kugelschalen unter sehr spitzen oder sehr stumpfen Winkeln, wird die Streubreite erhöht.

Die Maßeinheit des Dilution of Precision beschreibt dabei den Kehrwert der Fläche (bzw. des Volumens), die von den Einheitsvektoren vom Bestimmungsort in Richtung der Referenzpositionen aufgespannt wird. Ein hoher DOP-Wert steht für eine große Streubreite in der Messung, wohingegen ein kleiner DOP-Wert eine geringe Streubreite kennzeichnet. [16]

In Abbildung 3.8 wird das Phänomen der Dilution of Precision an einem Beispiel visualisiert. Die Entfernung wird zwischen einem Tag und zwei ortsfesten Anchors an den Positionen  $L_1$  und  $L_2$  gemessen. Der gesuchte Ort befindet sich in einem der Schnittpunkte auf der kreisförmigen Linie um die Referenzknoten mit dem Radius der jeweiligen Distanzen.

Die Distanzmessungen sind mit einer gewissen Messunsicherheit behaftet. In Abbildung 3.8 ist die Unsicherheit durch den Bereich zwischen den gestrichelten Linien dargestellt.

In Abbildung 3.8a liegen die beiden Referenzpunkte  $L_1$  und  $L_2$  weit auseinander und die Entfernungskreise treffen orthogonal aufeinander, sodass sich ein relativ kleiner Bereich ergibt, in dem der Tag mit einiger Sicherheit liegen muss.

In Abbildung 3.8b schneiden sich die Richtungsvektoren der Referenzpositionen in einem spitzen Winkel, was zu einer wesentlich höheren Unsicherheit führt. Obwohl die reine Messunsicherheit gleich geblieben ist, ist die Positionsunsicherheit erheblich größer.

Im Allgemeinen gilt, je mehr Referenzknoten für die Positionsbestimmung verwendet werden, desto kleiner sind die DOP-Werte und damit auch die erwartbaren Streubreiten der Positionsbestimmung. [16]



(a) Die beiden Entfernungskreise von  $L_1$  und  $L_2$  treffen orthogonal aufeinander. Die Positionsunsicherheit (grau gekennzeichnete Fläche) ist gering.

(b) Die Position des Anchors  $L_2$  ist näher an die Position des Anchors  $L_1$  herangerückt. Die Positionsunsicherheit (grau gekennzeichnete Fläche) ist erheblich größer.

Abbildung 3.8: Dilution of Precision

Quelle: In Anlehnung an [16]

In GPS-Systemen können hohe DOP-Werte in besonderen Umgebungen auftreten. In z. B. Häuser- und Straßenschluchten oder in stark bewaldeten Gebieten kann ein GPS-Empfänger aufgrund der Hindernisse unter Umständen nicht mit allen verfügbaren Satelliten kommunizieren. Werden nur GPS-Signale aus einem nur eingeschränkten Sichtfeld empfangen, sind die DOP-Werte groß und die Positionsgenauigkeit entsprechend ungenau. [16]

## 4 Ultra-wideband Lokalisationssystem

In diesem Kapitel ist das entwickelte UWB-Lokalisationssystem hardware- und softwareseitig dokumentiert. Unter Anwendung der Grundlagen aus den vorherigen Kapiteln wird insbesondere das eingesetzte Verfahren und Nachrichtenprotokoll zur Positionsbestimmung beschrieben.

### 4.1 Hardwareplattform

Für das Lokalisationssystem wird eine an der HAW Hamburg entwickelte Hardwareplattform verwendet. Das eingesetzte Hardwaredesign wurde im Rahmen der Bachelorarbeit von Herrn Markus Kasten entworfen und wird für diese Arbeit weiter verwendet. [14]

Die Abbildung 4.1 zeigt die Leiterplatte eines UWB-Beacons. Der vollständige Schaltplan befindet sich in Anhang A. Die Platine ist unter anderem mit einem *Decawave DWM1000* Ultra-wideband Transceiver und einer *Bosch BMI270* Inertial Measurement Unit (IMU) bestückt, die über einen Serial Peripheral Interface (SPI) Bus an einen Mikrocontroller der *STM32 L0*-Serie angebunden sind.

Im Folgenden werden die wichtigsten Eigenschaften und die Aufgaben dieser Komponenten vorgestellt:

**Decawave DWM1000** Der *Decawave DWM1000* Ultra-wideband Transceiver ist das Kernbauteil für die UWB-Kommunikation. Das Modul stellt einen vollständig IEEE 802.15.4-2011 kompatiblen Ultra-wideband Sender/Empfänger dar. [10]

Das Modul bietet Unterstützung für Echtzeit-Lokalisationssysteme, die mit Time of Flight (ToF) Messungen nach dem ToA oder TDoA Verfahren arbeiten. Dazu kann das Modul die Zeitpunkte des Nachrichtenempfangs bzw. des Nachrichtenversands mit einer zeitlichen Auflösung von 15,65 ps erfassen. [11]

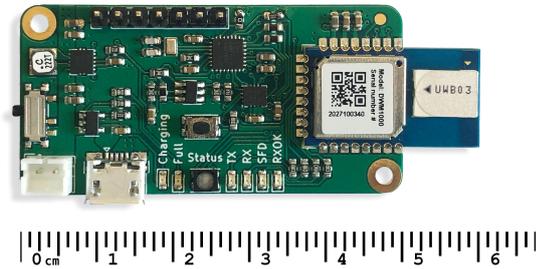


Abbildung 4.1: UWB-Beacon

Quelle: Eigene Darstellung

Aus dieser Zeitauflösung und der konstanten Lichtgeschwindigkeit in Luft ergibt sich für die minimal messbare Distanz eine theoretische Strecke  $s$  von

$$\begin{aligned} s &= c \cdot 15,65 \text{ ps} \\ &= 299\,702\,547 \text{ m s}^{-1} \cdot 15,65 \cdot 10^{-12} \text{ s} \\ &= 4,69 \cdot 10^{-3} \text{ m} = 4,69 \text{ mm} . \end{aligned}$$

Der Hersteller gibt die Genauigkeit in Distanzmessanwendungen hingegen realistischer mit  $\pm 10 \text{ cm}$  an. [10]

**Bosch BMI270** Die *Bosch BMI270* Inertial Measurement Unit (IMU) ist eine Kombination eines Beschleunigungssensors und eines Drehratensensors. Die Inertialsensoren können translatorische Bewegungen und rotierende Bewegungen in allen jeweils drei Freiheitsgeraden messen. [6]

Die Beschleunigung und Winkelgeschwindigkeit sind hilfreich für eine Sensordatenfusion, bei der die Ermittlung der Position durch diese Inertialmessungen unterstützt werden kann.

In dieser Arbeit wurden die Messwerte der IMU jedoch nicht eingesetzt. Dennoch wurden die Sensorwerte in allen Versuchen aufgezeichnet und stehen für spätere Nutzungen und Verbesserungen zur Verfügung.

**STM32 L0** Die *STM32 L0*-Serie ist eine Familie von Mikrocontrollern des Herstellers *STMicroelectronics* basierend auf einem 32-Bit ARM<sup>®</sup> Cortex<sup>®</sup>-M0+ Kern. Die Serie zeichnet sich besonders durch ihren geringen Energiebedarf aus.

Der eingesetzte Mikrocontroller ist mit 32 kB Flash-Speicher und 8 kB Arbeitsspeicher (RAM) ausgestattet und wird mit einer Taktfrequenz von 32 MHz betrieben. [22]

Der Mikrocontroller koordiniert im Wesentlichen die UWB-Kommunikation zur Entfernungsmessung, berechnet die Entfernungen und übermittelt die Messwerte über eine Universal Asynchronous Receiver Transmitter (UART) Schnittstelle an eine externe Lokalisierungseinheit zur Berechnung der Position. Die entsprechende Software ist in der Programmiersprache C implementiert.

Der UWB-Beacon kann sowohl als UWB-Anchor als auch als UWB-Tag eingesetzt werden. Die Funktionsweise wird durch die entsprechende Softwarekonfiguration bestimmt.

## 4.2 Protokoll zur Entfernungsmessung

Zur Positionsbestimmung wurde ein zweistufiges Verfahren ausgewählt. Zunächst werden die Entfernungen zwischen dem Tag und den verfügbaren Anchors gemessen. Die Entfernungsmessung basiert auf dem ToA Signalparameter und wird mit dem DS-TWR Protokoll umgesetzt. Aufgrund des geringen Fehlers durch zwei unabhängige Laufzeitmessungen ist dieses Verfahren von den in Abschnitt 3.1 vorgestellten Methoden am geeignetsten.

TDoA lässt sich mit der gegebenen Hardware hingegen nicht ideal umsetzen, da eine gemeinsame drahtlose Taktsynchronisation aller Anchors nur mit großem Protokollaufwand zu realisieren ist. AoA scheidet aufgrund der Notwendigkeit eines Antennenarray (zweites zusätzliches Antennenelement) aus.

In Abbildung 4.2 ist ein Sequenzdiagramm dargestellt, das die Umsetzung des DS-TWR Protokolls zur Entfernungsmessung und den Austausch der Zeitstempel anhand eines einfachen Beispiels mit einem Tag und einem Anchor zeigt.

Die drei Nachrichten des Protokolls sind als Poll Message, Response Message und Final Message wiederzufinden. Nach dem Nachrichtenaustausch wird die Entfernung auf dem Tag berechnet.

Zur Synchronisation wurde das Protokoll zudem um eine zusätzlich Blink Message erweitert. Genauere Informationen hierzu sind im nachfolgenden Abschnitt 4.4 zum Thema Zeitmultiplex zusammengestellt.

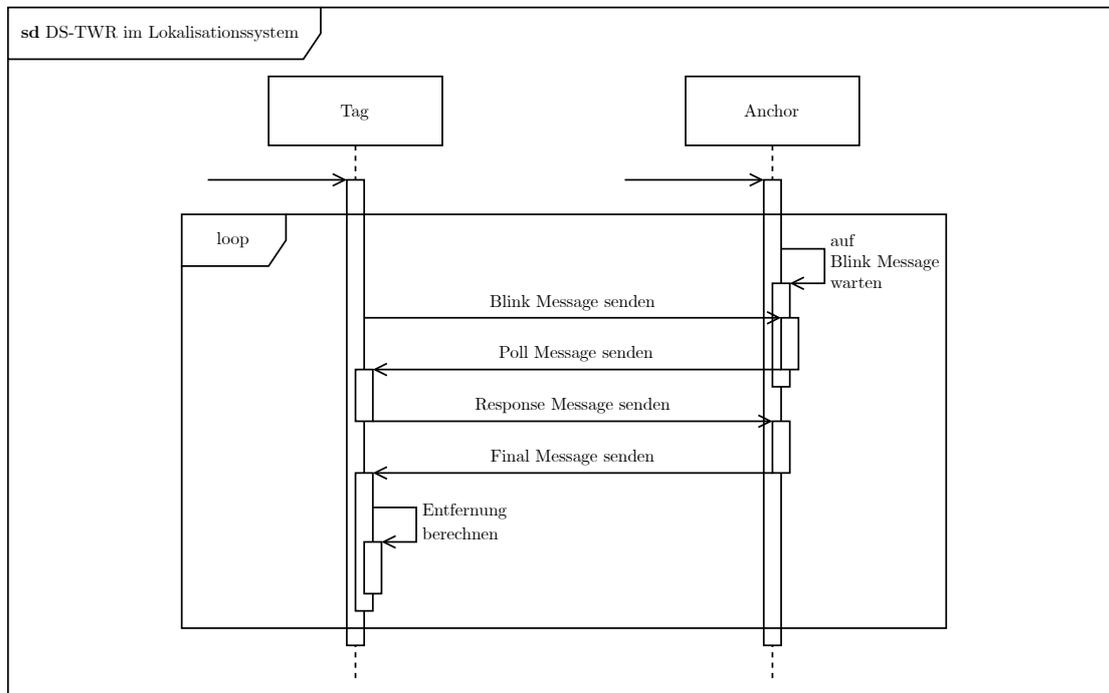


Abbildung 4.2: Sequenzdiagramm: DS-TWR im UWB-Lokalisationssystem

Quelle: Eigene Darstellung

Wie das Sequenzdiagramm (vgl. Abbildung 4.2) zeigt, kommen vier verschiedene Nachrichtentypen zum Einsatz. Die detaillierten Formate dieser Nachrichten sind im Folgenden dokumentiert.

Der Aufbau folgt den Konventionen der IEEE 802.15.4 Nachrichtenkodierung. Es handelt sich allerdings nur teilweise um standardisierte Nachrichten. Für weitere Einzelheiten sei auf die internationale Norm ISO/IEC 24730-62 verwiesen, die standardisierte Nachrichtenformate zur Verwendung in Echtzeit-Lokalisationssystemen auf der Grundlage von IEEE 802.15.4 enthält. [9]

Das allgemeine Nachrichtenformat basiert auf der IEEE 802.15.4 Standardkodierung. Dieses Schema wird in Abbildung 4.3 präsentiert. Die Tabelle 4.1 listet und beschreibt die einzelnen Felder. [9]

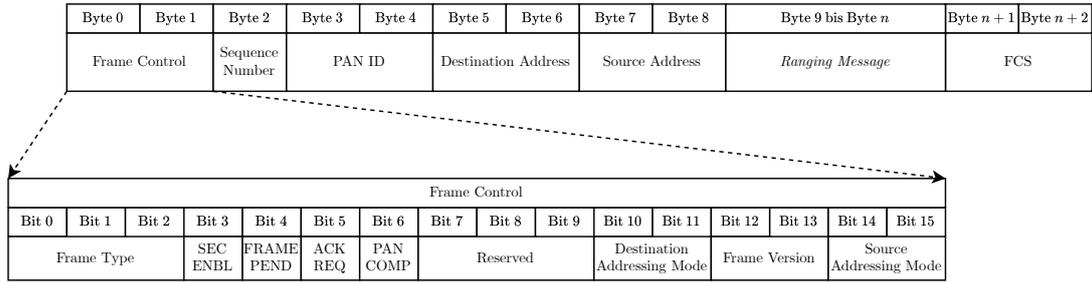


Abbildung 4.3: Allgemeines Nachrichtenformat nach IEEE 802.15.4

Quelle: In Anlehnung an [9]

Byte 9
Function Code

(a) Blink Message

Byte 9
Function Code

(b) Poll Message

Byte 9	Byte 10	Byte 11	Byte 12
Function Code	Activity Code	Activity Parameter	

(c) Response Message

Byte 9	Byte 10	Byte 11	Byte 12	Byte 13	Byte 14	Byte 15	Byte 16	Byte 17	Byte 18	Byte 19	Byte 20	Byte 21
Function Code	Poll TX Timestamp				Response RX Timestamp				Final TX Timestamp			

(d) Final Message

Abbildung 4.4: Kodierungen der Nachrichten zur Entfernungsmessung

Quelle: In Anlehnung an [9]

Tabelle 4.1: Aufbau des IEEE 802.15.4 Nachrichtenformats

Byte #	Wert	Feld	Beschreibung
0-1	0x8841	Frame Control	Das Frame Control Feld enthält Statusinformationen über den Nachrichtentyp, die Fragmentierung und andere protokollrelevante Daten. Der Wert 0x8841 legt eine geeignete Definition für den vorliegenden Anwendungsfall fest.
2	-	Sequence Number	Die Sequenznummer gewährleistet eine reihenfolgegesicherte Nachrichtenfolge und das Erkennen von Duplikaten. Der Wert wird mit jeder gesendeten Nachricht inkrementiert.
3-4	0x1DEA	PAN ID	Die PAN ID wird zu eindeutigen Identifizierung des Personal Area Networks (PAN) verwendet. Die PAN ID ist für alle Knoten desselben Lokalisationssystems gleich. Der verwendete Wert 0x1DEA ist willkürlich gewählt.
5-6	-	Destination Address	Mit der Zieladresse wird der Empfänger der Nachricht eindeutig identifiziert.
7-8	-	Source Address	Mit der Quelladresse wird der Sender der Nachricht eindeutig identifiziert.
9- $n$	-	Ranging Message	Der Inhalt dieses Nachrichtenbereichs hängt davon ab, um welche der vier Nachrichten zur Entfernungsmessung es sich handelt.
$n + 1$ - $n + 2$	-	FCS	Die Frame Check Sequence (FCS) dient der Fehlererkennung und enthält dazu eine CRC-Prüfsequenz. Der CRC wird automatisch durch den UWB-Transceiver ausgeführt und die Prüfsequenz an die übertragende Nachricht angehängt.

Mit dem Frame Control können protokollrelevante Daten wie der Nachrichtentyp oder die Adressierungsart festgelegt werden. Die zwei Bytes des Frame Control Feldes sind für die Anwendung konstant belegt. Mit dieser Konfiguration wird im Wesentlichen die Ver-

wendung von 16-Bit Quell- und Zieladressen und die Verwendung einer einzelnen 16-Bit PAN ID festgelegt. Die Tabelle 4.2 enthält und beschreibt die gesamten spezifizierbaren Optionen des Frame Control Feldes. [9]

Tabelle 4.2: Aufbau des IEEE 802.15.4 Frame Control Feldes

Bit #	Wert	Feld	Beschreibung
0-2	0b100	Frame Type	Das Feld mit dem Wert 0b100 spezifiziert den Typ dieses Datenframes als allgemeine Daten-Nachricht.
3	0b0	Security Enabled	Dieses Feld ist auf den Wert 0b0 gesetzt, um zu signalisieren, dass die Nachricht nicht geschützt ist.
4	0b0	Frame Pending	Der Wert 0b0 signalisiert, dass aktuell keine weiteren Daten für den Empfänger vorliegen.
5	0b0	ACK Request	Der Wert 0b0 legt fest, dass keine Bestätigungsmeldung vom Empfänger benötigt wird.
6	0b1	PAN ID Compression	Dieses Feld mit dem Wert 0b1 legt fest, dass die Quell- und Zieladresse aus dem selben PAN stammen. Hierdurch ist die PAN ID nur einmal in der Nachricht enthalten.
7-9	-	Reserved	Nicht verwendet.
10-11	0b01	Destination Addressing Mode	Dieses Feld mit dem Wert 0b01 legt die Verwendung von 16-Bit Zieladressen fest.
12-13	0b00	Frame Version	Das Feld ist auf den Wert 0b00 gesetzt, um eine IEEE 802.15.4-2003 kompatible Nachricht zu kennzeichnen.
14-15	0b01	Source Addressing Mode	Dieses Feld mit dem Wert 0b01 legt die Verwendung von 16-Bit Quelladressen fest.

Der Inhalt des Ranging Message Feldes hängt davon ab, um welchen der vier Nachrichtentypen es sich handelt. Diese Typen sind in Abbildung 4.4 dargestellt und in den Abschnitten 4.2.1 bis 4.2.4 beschrieben. In diesen Abschnitten wird nur der jeweilige Teil des Ranging Message Feldes gezeigt und erläutert. Die Daten werden in das allgemeine

Nachrichtenformat aus Abbildung 4.3 eingebettet, um jeweils die vollständige Nachricht zu bilden. [9]

### 4.2.1 Blink Message

Die Blink Message dient zur Synchronisierung der Anchors. Die Tabelle 4.3 listet die individuellen Felder der Blink Message auf und beschreibt ihre Bedeutungen. [9]

Tabelle 4.3: Felder in der Blink Message

Byte #	Wert	Feld	Beschreibung
9	0x60	Function Code	Das Byte mit dem Wert 0x60 kennzeichnet die Nachricht als Blink Message.

### 4.2.2 Poll Message

Die Poll Message wird vom UWB-Anchor zur Initiierung des Entfernungsmessvorgangs gesendet. Die Tabelle 4.4 listet die individuellen Felder der Poll Message auf und beschreibt ihre Bedeutungen. [9]

Tabelle 4.4: Felder in der Poll Message

Byte #	Wert	Feld	Beschreibung
9	0x21	Function Code	Das Byte mit dem Wert 0x21 kennzeichnet die Nachricht als Poll Message.

### 4.2.3 Response Message

Die Response Message wird vom UWB-Tag als Antwort auf die Poll Message gesendet. Im DS-TWR Protokollablauf beendet diese Nachricht die erste Messung der Paketumlaufzeit und startet den zweiten Messvorgang. Die Tabelle 4.5 listet die individuellen Felder der Response Message auf und beschreibt ihre Bedeutungen. [9]

Tabelle 4.5: Felder in der Response Message

Byte #	Wert	Feld	Beschreibung
9	0x10	Function Code	Das Byte mit dem Wert 0x10 kennzeichnet die Nachricht als Response Message.
10	0x02	Activity Code	Der Activity Code 0x02 teilt dem Anchor mit, dass mit dem Entfernungsmessaustausch fortgefahren werden soll.
11-12	0x00	Activity Parameter	Dieses Feld ist für optionale Parameter für den Activity Code reserviert. Für den genutzten Activity Code 0x02 sind keine Parameter vorgesehen.

#### 4.2.4 Final Message

Die Final Message wird vom UWB-Anchor nach Erhalt der Response Message an den UWB-Tag gesendet und beendet den gesamten Entfernungsmessvorgang. Die Nachricht ist 24 Bytes lang und enthält alle nötigen Zeitstempel, die auf dem UWB-Tag zur Errechnung der Distanz zwischen Anchor und Tag erforderlich sind. Die Tabelle 4.6 listet die individuellen Felder der Final Message auf und beschreibt ihre Bedeutungen. [9]

Tabelle 4.6: Felder in der Final Message

Byte #	Wert	Feld	Beschreibung
9	0x23	Function Code	Das Byte mit dem Wert 0x23 kennzeichnet die Nachricht als Final Message.
10-13	-	Poll TX Timestamp	Dieses Feld beinhaltet den genauen Zeitstempel für den Sendezeitpunkt der Poll Message.
14-17	-	Response RX Timestamp	Dieses Feld beinhaltet den genauen Zeitstempel für den Empfangszeitpunkt der Response Message.
18-21	-	Final TX Timestamp	Dieses Feld beinhaltet den genauen Zeitstempel für den Sendezeitpunkt der Final Message. Der Wert wird auf dem Anchor vorausberechnet.

### 4.3 Technische Daten

Die Kommunikation des UWB-Lokalisationssystems findet auf UWB-Kanal 2 statt, was einer Mittenfrequenz von 3993,6 MHz entspricht. Bei einer Bandbreite von 449,2 MHz folgt daraus ein Frequenzbereich von 3774 MHz – 4243,2 MHz. Die Datenrate beträgt 110 kHz. [11]

Die Preamble und der SFD der UWB-Nachrichtenpakete sind auf eine Länge von 1024 Symbolen bzw. 64 Symbolen festgelegt. Durch diese Wahl einer langen Preamble und einer geringen Datenrate erhöht sich die Robustheit und Reichweite des Systems. Gleichzeitig steigt jedoch die Nachrichtenlänge, wodurch sich die maximal mögliche Anzahl der UWB-Beacons und die Messrate des Systems reduziert. Dieser Zusammenhang wird im folgenden Abschnitt 4.4 verdeutlicht. Die durchschnittliche Nachrichtenlänge beträgt bei der zuvor beschriebenen Parametrisierung etwa 3 ms.

Mit einer Frequenz von 10 Hz bestimmt der Tag seine Entfernung zu den verfügbaren Anchors. In einem UWB-Lokalisationssystem mit einem Tag und fünf Anchors erfolgen somit 50 einzelne Entfernungsmessungen pro Sekunde.

Im Test konnte eine maximale Reichweite von ungefähr 30 m zwischen Sender und Empfänger ermittelt werden.

### 4.4 Zeitmultiplexverfahren

Um die interferenzfreie Kommunikation mit mehreren UWB-Beacons zu unterstützen, wird ein Time Division Multiple Access (TDMA) Zeitmultiplexverfahren eingesetzt.

Das Gegenteil zu einem Zeitmultiplexverfahren stellt die triviale ALOHA-Methode dar, mit der es einem Teilnehmer zu jedem möglichen Zeitpunkt erlaubt ist, verfügbare Daten auf den gemeinsamen Kommunikationskanal zu senden. Dabei existiert stets die Gefahr, dass zwei oder mehrere Datenpakete kollidieren, wodurch die Informationen aller Nachrichten verloren gehen. Das Resultat ist die Abnahme der Systemperformance. Statistische Überlegungen zeigen, dass mit der ALOHA-Methode die Übertragungskapazität eines Kommunikationskanals nur zu maximal 18 % ausgenutzt werden kann. [24]

Für komplexere Protokolle oder Echtzeitanforderungen, wie im Falle des UWB-Lokalisationssystems, ist die ALOHA-Methode daher nicht geeignet.

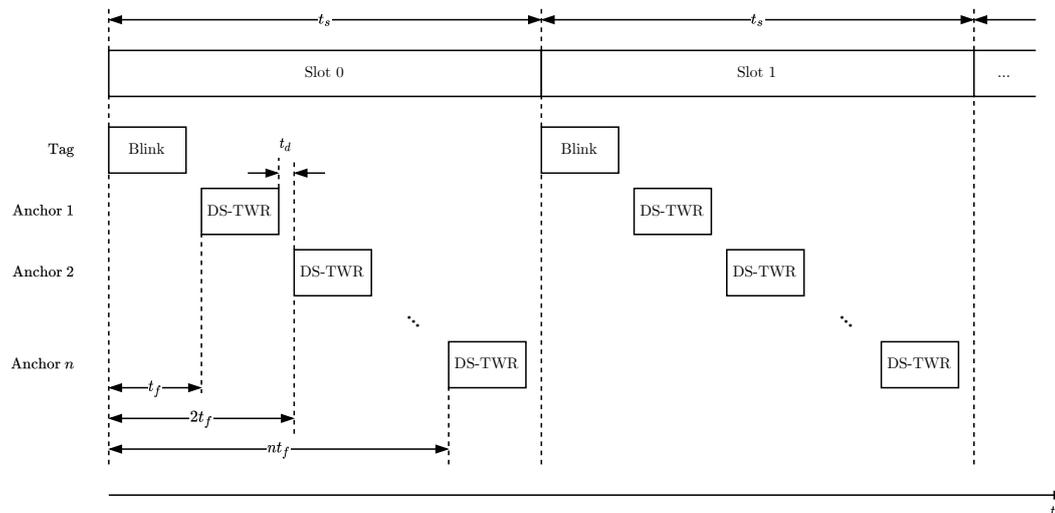


Abbildung 4.5: Zeitprofil zur gemeinsamen Nutzung des Kommunikationskanals

Quelle: In Anlehnung an [9]

Ein Zeitmultiplexverfahren gewährleistet hingegen die koordinierte gemeinsame Nutzung eines Kommunikationskanals, indem jeder Sender durch einen Multiplexer einen festen Zeitabschnitt zum Senden seiner Daten zugeteilt bekommt. Das eingesetzte Verfahren ist in Abbildung 4.5 dargestellt.

Der Tag übernimmt die Rolle des Multiplexers, indem er eine Blink Message an alle Teilnehmer des Systems sendet, um den Beginn eines Zeitslots anzukündigen. Der erste Anchor startet daraufhin nach einer zeitlichen Verzögerung von  $t_f$  mit der DS-TWR Entfernungsmessung. Der zweite Anchor wartet eine Zeit von  $2t_f$ , bevor er seine Distanzmessung mit dem Tag initiiert. In dieser Zeit muss die Kommunikation des vorherigen Anchors abgeschlossen sein. Der Ablauf setzt sich nach dem gleichen Schema fort, bis die Zeit  $t_s$  des Zeitslots abgelaufen ist und der Tag mit einer neuen Blink Message den Beginn eines weiteren Zeitslots ankündigt. Auf diese Weise wird jedem Anchor zu einem bestimmten Zeitpunkt eine exklusive Kommunikation mit dem Tag zugeteilt.

In der gegenwärtigen Implementation mit einer Messfrequenz von 10 Hz beträgt die Zeit  $t_s$  für einen Slot 100 ms. Die Verzögerungszeit  $t_f$  ist mit 10 ms gewählt, sodass jeder Slot in zehn gleichlange Zeitabschnitte zerfällt. Folglich unterstützt das UWB-Lokalisationssystem eine Gesamtzahl von neun UWB-Anchors. Zukünftig ließe sich diese Zahl durch die Verwendung einer höheren Datenrate oder einer kürzeren Preamble

steigern, was zu kürzeren Nachrichtenlängen führt, sodass die Verzögerung  $t_f$  verkürzt werden kann.

### 4.5 Lokalisierungseinheit

Abhängig davon, wo die Positionsbestimmung vorgenommen wird, lassen sich die Lokalisierungsverfahren in endgerätebasierte oder netzwerkbasierte Systeme einteilen. Bei der endgerätebasierten Lokalisierung empfängt das jeweilige Ziel die Signale von den Referenzknoten und errechnet seinen eigenen Standort (Selbstpositionierung). Diese Methode wird üblicherweise im militärischen Bereich oder im Bereich der öffentlichen Sicherheit eingesetzt. GPS-Systeme fallen ebenfalls in diese Kategorie.

Bei der netzwerkbasierten Lokalisierung leiten die Teilnehmer die empfangenen Signalinformationen, wie z. B. RSSI, ToA, oder AoA an eine zentrale Lokalisierungseinheit weiter, wo der Standort des Ziels ermittelt wird (Remote-Positioning). [20]

In dem vorgestellten UWB-Lokalisationssystem wird die Position in einem netzwerk-basierten Verfahren bestimmt. Die zentrale Lokalisierungseinheit ist in der Programmiersprache Python implementiert. Durch die Wahl von ToA Signalinformationen zur Distanzmessung zwischen Tag und Anchor erfolgt die Positionsberechnung mittels der Trilateration. Das Trilaterationsproblem wird dabei durch den Sequential Least Squares Programming (SLSQP) Optimierungsalgorithmus gelöst. Zudem wurden mit der Lokalisierungseinheit mehrere Filter umgesetzt, die die Positionierungsgenauigkeit verbessern sollen. Diese Filter werden im anschließenden Kapitel ausführlich behandelt. Ein Bildschirmfoto der Anwendung wird in Abbildung 4.6 gezeigt.

### 4.6 Praxistest

Während der Entwicklung wurde das UWB-Lokalisationssystem mehrfach in Praxissituationen getestet. Durch die Kooperation mit anderen Projekten und Forschungsgruppen der HAW Hamburg (z. B. *Shared Guide Dog 4.0* [12]) erfolgten die Tests durchaus unter realitätsnahen Bedingungen mit unterschiedlichen Anforderungen, sowohl im Innen- als auch im Außenbereich.

Das Lokalisationssystem wurde zudem sehr umfangreich mit einem optischen AR-Trackingsystem getestet. Die HAW Hamburg verfügt mit dem Creative Space for Technical

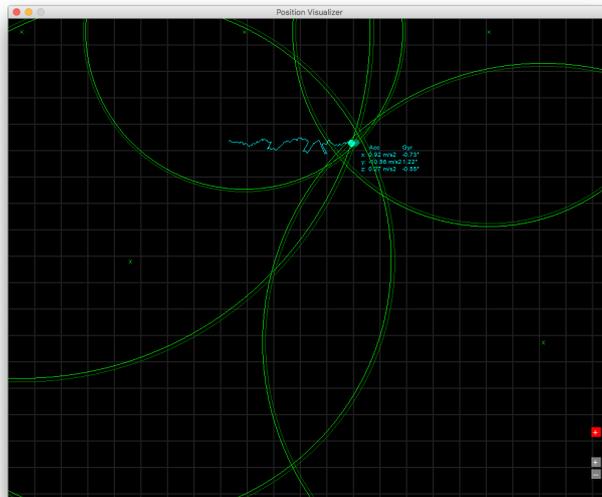


Abbildung 4.6: Bildschirmfoto der zentralen Lokalisierungseinheit.  
Die Entfernungskreise sind grün und das verfolgte Ziel ist türkis eingezeichnet.

Quelle: Eigene Darstellung



Abbildung 4.7: Praxistest des UWB-Lokalisationssystems mit einem AR-Trackingsystem im CSTI

Quelle: Eigene Darstellung

Innovations (CSTI) über das AR-Trackingsystem *ARTTRACK5* des Herstellers *ART*. Das Trackingsystem des CSTI arbeitet mit acht Infrarotkameras, die die Pose von speziell infrarotreflektierenden Tracking-Targets erfassen können. Die Auflösung des System liegt bei 0,1 mm mit einer Messfrequenz von 60 Hz. Damit ist das AR-Trackingsystem geeignet, um präzise Ground Truth Informationen als Referenz zum UWB-Lokalisationssystem zu ermitteln. [2]

Die erhobenen Daten und die erzielte Genauigkeit werden im nachfolgenden Kapitel ausführlicher diskutiert. Der Test im CSTI ist in Abbildung 4.7 dokumentiert.

### 4.7 Robustheit

Durch den Einsatz von UWB-Signalen ist die drahtlose Kommunikation weitestgehend unempfindlich gegenüber äußere Störungen. Die Parametrisierung der UWB-Signale (lange Preamble, geringe Datenrate) wurde besonders in Hinblick auf eine hohe Robustheit vorgenommen.

Das Lokalisationssystem ist ebenfalls robust gegenüber Änderungen in der Systemstruktur. Allerdings ist zu berücksichtigen, dass mindestens drei UWB-Anchors in direkter Reichweite sein müssen, um eine Positionsbestimmung im zweidimensionalen Raum auszuführen. Darüber hinaus zusätzliche UWB-Anchors können gemäß der Dilution of Precision die Streubreite verringern und die Genauigkeit erhöhen, sind aber nicht zwingend erforderlich.

Für den Einsatz in Produktivumgebungen ist es dennoch empfehlenswert auf weitere Sensorik und Möglichkeiten zur Positionsermittlung zurückgreifen zu können, um beispielsweise Ausfälle zu kompensieren. Zudem können spezialisierte Messsysteme eingesetzt werden, wenn in bestimmtem (Ziel-)Bereichen sehr hohe Genauigkeiten oder Hinderniserkennungen benötigt werden. Je nach Anwendungsfall und Anforderungen eignen sich z. B. GPS-, Lidar- oder Radar-Systeme sowie Tiefenkameras.

In der Robotik sind zudem Simultaneous Localization and Mapping (SLAM) Verfahren bekannt, bei denen ein mobiler Roboter selbstständig eine Karte der Umgebung erstellt und seine eigene Position innerhalb dieser Karte bestimmt. [21]

Diese Sensoren und Verfahren werden in dieser Arbeit jedoch nicht weiter betrachtet.

# 5 Verbesserung der Lokalisierungsgenauigkeit

In den vergangenen Jahrzehnten wurden bereits viele Möglichkeiten beschrieben und Ansätze entwickelt, um die Zuverlässigkeit und Genauigkeit der GPS-Navigation zu verbessern. Aufgrund einer ähnlichen Domäne können viele der Ansätze auch auf die Verbesserung der Positionsbestimmung in UWB-Systemen übertragen werden.

In GPS-Anwendungen haben sich besonders Methoden, die auf Kalman-Filtern basieren, etabliert. Zudem sind in der Robotik Particle-Filter bekannt, die stark verbreitet in Kartierungs- und Navigationsaufgaben eingesetzt werden. [27, 21]

In dieser Arbeit wird passend zur neuartigen UWB-Positionsbestimmung ein ebenfalls moderner Ansatz zur Verbesserung der Lokalisierungsgenauigkeit erprobt. Mithilfe von Reinforcement Learning soll eine optimale Strategie zur Korrektur der Positionsmessungen erlernt werden.

Reinforcement Learning ist ein maschinelles Lernverfahren, bei dem ein Agent durch die Interaktion mit seiner Umgebung lernt, ein bestimmtes Ziel auf optimalen Weg zu erreichen. Dazu probiert der Agent verschiedene Aktionen aus und nimmt damit Einfluss auf die Umgebung. Von der Umgebung bekommt der Agent eine Bewertung seiner Aktionen. Durch passend gewählte Aktionen maximiert der Agent diese Bewertungen. [23]

In den nachfolgenden Abschnitten werden ein RL-Modell, ein Kalman-Filter und ein Particle-Filter zur Verbesserung der Lokalisierungsgenauigkeit beschrieben und die Leistungen in einem anschließenden Benchmark miteinander verglichen und bewertet.

## 5.1 Fehler

Wie gut ein Lokalisationssystem funktioniert, wird üblicherweise mit der erzielten Genauigkeit gemessen. Die Lokalisierungsgenauigkeit ist definiert als der Abstand zwischen der

bestimmten Position und der tatsächlichen Position des Zielobjektes (z. B.  $\pm 1$  m Genauigkeit). Zusätzlich ist es wünschenswert, dass eine bestimmte Lokalisierungsgenauigkeit mit hoher Wahrscheinlichkeit erreicht wird. Die Präzision definiert den Prozentsatz, mit dem eine bestimmte Genauigkeit (oder besser) erreicht wird (z. B. 95 % Präzision). [20]

Der Fehler, der die Genauigkeit und die Präzision des UWB-Lokalisationssystems beeinträchtigt, stammt aus mehreren Quellen. Unsicherheiten in der Angabe der bekannten Positionen der UWB-Anchors führen zu konstanten Messabweichungen. Darüber hinaus wirken sich unter anderem Temperaturänderungen und Phänomene wie die Mehrwegeausbreitung und Abschattung als Rauschen auf die Messungen aus. Infolge der Nutzung des ToA Ansatzes zur Entfernungsmessung ist das Rauschen unabhängig zwischen den verschiedenen Teilnehmern. [26]

## 5.2 Reinforcement Learning

Der vorgestellte RL-Agent verbessert die Genauigkeit von  $x$ - und  $y$ -Koordinaten der Positionen, die von dem UWB-Lokalisationssystem bestimmt werden, durch geeignete Korrekturmaßnahmen.

Im Gegensatz zu anderen hardware- oder softwarebasierten Methoden benötigt der Agent dazu keine zusätzlichen Komponenten (z. B. zusätzliche Sensoren) und arbeitet intern nicht mit starren Annahmen über das Messrauschen des Lokalisationssystems oder mit Bewegungsmodellen. Aus diesen Gründen ist besonders gegenüber dem Kalman-Filter ein Vorteil zu erwarten, da klassische Kalman-Filter auf einer statischen Systemmodellierung und festen Annahmen über die Rauschparameter des Systems beruhen. Durch Reinforcement Learning kann stattdessen eine optimal auf den Problemfall zugeschnittene Strategie (engl. policy) entwickelt werden, womit die Positionsmessungen korrigiert werden. [27]

Als Vorbild für den RL-Agenten dient die Arbeit von Zhang und Masoud [27], die in ihrer Veröffentlichung „*Increasing GPS Localization Accuracy With Reinforcement Learning*“ ein RL-Modell zur Korrektur von GPS-Positionsmessungen einführen. Der vorgestellte Ansatz ist besonders geeignet, da die Arbeitsweise vollständig vom eingesetzten Lokalisationssystem abgekoppelt ist. In Abbildung 5.1 wird anhand des dargestellten Informationsflusses deutlich, dass sich das RL-Modell ähnlich eines Filters verhält und nur

die Ausgänge des Lokalisationssystems als Eingabe verwendet. Aufgrund dieser Eigenschaft ist das Verfahren nicht auf GPS-Systeme beschränkt und kann ebenso in UWB-Anwendungen genutzt werden. [27]

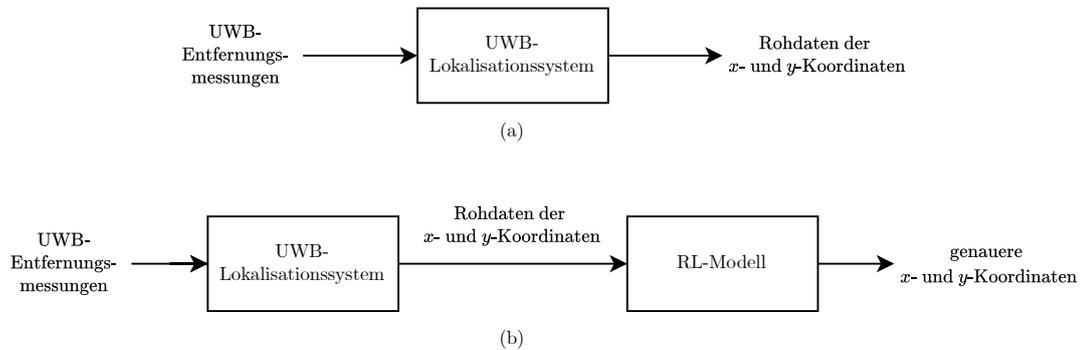


Abbildung 5.1: Vergleich des Datenflusses einer typischen Positionsbestimmung (a) und einer Positionsbestimmung mit nachgeschaltetem RL-Modell (b)

Quelle: In Anlehnung an [27]

Aus der Perspektive des Reinforcement Learning ist insbesondere der eingesetzte Mechanismus zur Ermittlung der Belohnung (engl. reward) interessant. Eine Herausforderung besteht darin, dass die Belohnung unabhängig von der tatsächlichen Position sein muss, da in einer praktischen Anwendungen kein Ground Truth als Referenz zur Verfügung steht. Die Lösung für dieses Problem ist ein abstrakter Belohnungsmechanismus, der nicht wie üblicherweise direkte Belohnungen erteilt, sondern sich auf ein Vertrauensmaß der Modellvorhersagen bezieht. [27]

Die Funktionsweise des verwendeten Belohnungsmechanismus ist in Abschnitt 5.2.3 genauer erläutert. Weiterhin werden in den nachfolgenden Absätzen die Details des RL-Ansatzes, einschließlich des Beobachtungsraums, des Aktionsraums, des RL-Algorithmus und der Trainingsarchitektur vorgestellt.

### 5.2.1 Beobachtungsraum

Das UWB-Lokalisationssystem meldet in regelmäßigen Zeitabständen eine aktuelle Position. Für das vorgestellte RL-Modell beschränkt sich eine Beobachtung nicht nur auf die zuletzt gemeldete Position, sondern ist vielmehr ein Vektor, der sowohl die zuletzt bestimmte Position als auch die letzte Vorhersagehistorie von korrigierten Positionen

enthält. Dieser Ansatz zur Erzeugung eines Beobachtungsvektors ermöglicht es dem Modell, Informationen über die zurückgelegte Trajektorie des UWB-Tags sowie die eigenen vergangenen Schätzungen zu berücksichtigen. [27]

Es sei  $q_t$  die vom UWB-Lokalisationssystem gemessene Position zum Zeitpunkt  $t$ . Der zugehörige Ground Truth  $g_t$  ist unbekannt. Daher kann das Problem als Partially Observable Markov Decision Process (POMDP) formuliert werden. In diesem POMDP wird  $p_t$  verwendet, um den Glaubenszustand (engl. belief state) des gemessenen Punktes  $q_t$  zu bezeichnen. [27]

Poupart und Boutilier [19] beschreiben, dass in einem POMDP ein optimaler Aktionsablauf gefunden werden kann, indem ein Markov Decision Process (MDP) mit vollständig beobachtbaren Glaubenszuständen angenommen wird. Die Glaubenszustände abstrahieren dabei die nur partiell beobachtbaren Zustände des POMDP, sodass anhand dieser Schätzungen ein MDP gebildet werden kann. [19, 27]

Weiterhin wird ein Beobachtungsvektor  $Z_t$  mit konstanter Größe  $n$  verwendet, um die historischen Modellschätzung der letzten  $n-1$  Positionen und das aktuelle  $q_t$  zu speichern; d. h.  $Z_t = \{p_{t-n-1}, \dots, p_{t-1}, q_t\}$ .

Mit  $S_t$  und  $b_t$  wird der interne Zustand bzw. der Glaubenszustand von  $Z_t$  bezeichnet. Bei einem Beobachtungsvektor der Größe  $n$  enthält der Vektor  $S_t$  zum Zeitpunkt  $t$  den entsprechenden Ground Truth dieser Punkte; d. h.  $S_t = \{g_{t-n}, \dots, g_t\}$ . Der Vektor  $b_t$  enthält die Vorhersagen der letzten  $n$  Punkte; d. h.  $b_t = \{p_{t-n}, \dots, p_t\}$ .

Die Vektoren  $Z_t$  und  $b_t$  unterscheiden sich also nur in ihrem letzten Element. Das letzte Element von  $Z_t$  ist  $q_t$ , während das letzte Element des Vektors  $b_t$  die Vorhersage  $p_t$  ist. [27]

Das Modell schätzt  $b_t$  auf der Grundlage von  $Z_t$ . Das resultierende POMDP-Diagramm ist in Abbildung 5.2 dargestellt, wobei die Aktionen mit  $a$  und die Belohnungen mit  $R$  bezeichnet werden. [27]

Ein Modell des entsprechenden RL-Agenten ist in Abbildung 5.3 dargestellt. Zur Zustandsschätzung werden die Beobachtungen, Aktionen und die Glaubenszustände als Eingabe verwendet, woraus der Zustandsschätzer einen angepassten Glaubenszustand als Ausgabe liefert. Die erlernte Policy des RL-Agenten bildet Glaubenszustände auf Aktionen ab. Die ausgeführte Aktion beeinflusst die Beobachtung und den Glaubenszustand des nächsten Zeitschritts. [27]

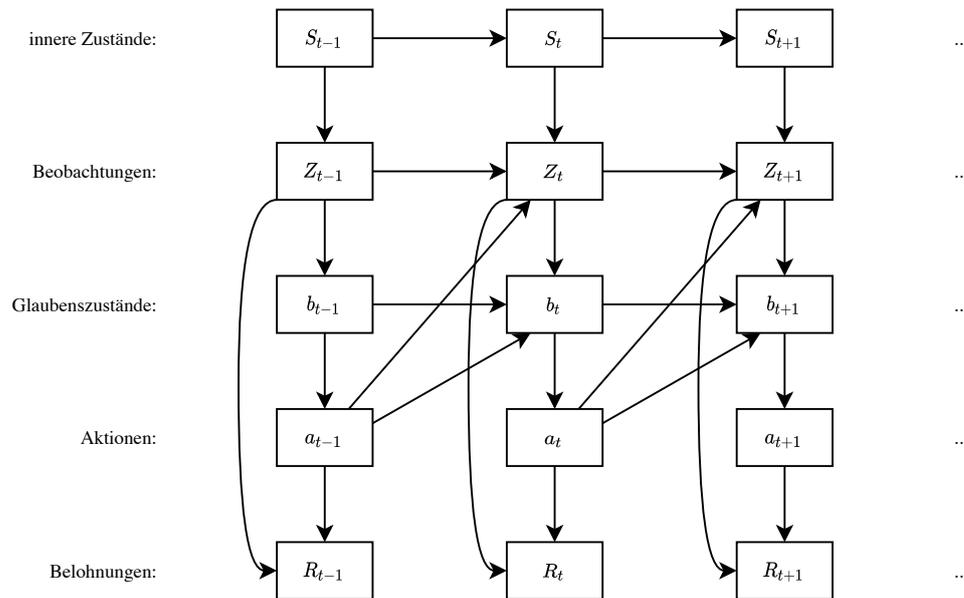


Abbildung 5.2: POMDP-Diagramm des RL-Modells

Quelle: In Anlehnung an [27]

Basierend auf der zuvor beschriebenen Definition des POMDP besteht das Ziel des RL-Agenten in jedem Zeitschritt  $t$  darin, die optimale Aktion zur Korrektur von  $q_t$  zu finden. Der Prozess schreitet kontinuierlich in einem Schiebefensterverfahren (engl. sliding window) voran. Sobald ein neues  $q_t$  aus dem UWB-Lokalisationssystem vorliegt, bewegt sich das Schiebefenster einen Schritt vorwärts. Der Beobachtungsvektor  $Z_t$  wird aktualisiert, wobei  $q_t$  das letzte Element darstellt und die letzten  $n-1$  Modellvorhersagen den restlichen Vektor auffüllen. Der Aktualisierungsprozess ist in Abbildung 5.4 dargestellt. [27]

Jedes Mal, wenn das UWB-Lokalisationssystem eine neue Position  $q_t$  meldet, wird der Beobachtungsvektor aktualisiert und das Modell trainiert. Wenn  $p_t$  bestimmt wird, wird diese Vorhersage in den Beobachtungsvektor verschoben, um  $q_t$  zu ersetzen, während die Zeit auf  $t+1$  erhöht wird. Das heißt, in jedem Trainingsschritt enthält der Beobachtungsraum des RL-Modells sowohl die aktuell gemessene Position als auch eine Reihe von historischen Positionsschätzungen. [27]

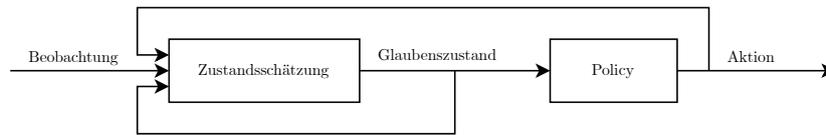


Abbildung 5.3: Modell des RL-Agenten

Quelle: In Anlehnung an [27]

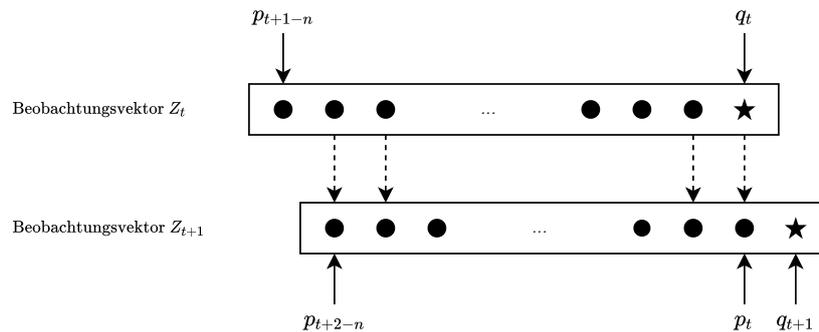


Abbildung 5.4: Aktualisierungsprozess der RL-Modelleingaben

Quelle: In Anlehnung an [27]

### 5.2.2 Aktionsraum

Eine Aktion des Agenten ist definiert als ein Vorgang zur Anpassung von  $x$ - und  $y$ -Koordinaten.

Zur Reduzierung der Komplexität werden die kontinuierlichen Werte der Koordinaten in kleine Schritte diskretisiert. Grundsätzlich ist eine Diskretisierung des Aktionsraums im Allgemeinen nicht zu empfehlen, da sich die Größe der Policy-Tabelle exponentiell erhöhen kann. Allerdings ist in niedrigdimensionalen Aktionsräumen, wie es bei diesem Problem der Fall ist, die damit verbundene Reduktion des Rechenaufwandes des Algorithmus vorteilhaft. [27]

Für den Agenten wurde der Aktionsraum nach folgendem Verfahren diskretisiert:

1. Es werden Operationseinheiten zur Anpassung der  $x$ - und  $y$ -Koordinaten definiert, die jeweils mit  $u_x$  und  $u_y$  bezeichnet sind, sowie eine Skalierungsmatrix  $M$ . Die Skalierungsmatrix  $M$  beinhaltet das Intervall  $M = [-5; 5]$ .
2. Die möglichen Operationsbereiche zur Korrektur der  $x$ - und  $y$ -Koordinaten sind mit  $L_x$  bzw.  $L_y$  bezeichnet. Dabei ist  $L_x = M \cdot u_x$  und  $L_y = M \cdot u_y$ .
3. Die durchführbaren Aktionen bilden eine diskrete Aktionsmenge  $A$ , indem die Operationen auf den  $x$ - und  $y$ -Koordinaten kombiniert werden. Eine diskrete Aktion  $a \in A$  kann definiert werden als  $a = (x, y) \in L_x \times L_y$ .
4. Die Operationen auf den  $x$ - und  $y$ -Koordinaten werden durch die Aktionsmenge  $A$  dargestellt, wobei  $A$  alle möglichen Aktionen enthält. Die Kardinalität der endlichen Aktionsmenge  $A$  ist  $11^2 = 121$ , basierend auf der entsprechend oben angegebenen Skalierungsmatrix  $M$ .

Die Durchführung von Aktionen aus Aktionsraum  $A$  auf einer gemessenen Position kann zu 121 möglichen Ausgaben führen. Wenn eine Position beispielsweise  $(x, y)$  ist, dann sind  $(x \pm mu_x, y \pm nu_y)$ ,  $m, n \in M$  mögliche Korrekturausgaben.

Durch die Wahl dieser Diskretisierung kann der Agent nur Korrekturen durchführen, die im Bereich der ursprünglich gemessenen Koordinaten liegen. Eine zuverlässige Korrektur ist also nicht möglich, wenn ein Punkt sehr extrem von der wirklichen Position abweicht. Der Korrekturbereich kann durch die Wahl unterschiedlicher Werte für  $u_x$  und  $u_y$  geändert werden. Diese Eigenschaft des Aktionsraums muss bei der Wahl von  $u_x$  und  $u_y$  berücksichtigt werden. Die geeigneten Werte für  $u_x$  und  $u_y$  werden somit durch die Genauigkeit des UWB-Lokalisationssystems vorgegeben. [27]

### 5.2.3 Belohnungsmechanismus

In der Robotik werden Konfidenzellipsen in SLAM Verfahren verwendet, um die Zuverlässigkeit der Lokalisierung zu messen. Ein SLAM Verfahren besteht darin, dass sich ein autonomer Roboter in einer Umgebung bewegt, in der es verschiedene Orientierungspunkte gibt. Das Ziel des Roboters ist es, eine Karte der Umgebung zu erstellen und sich gleichzeitig so genau wie möglich in der Karte zu lokalisieren. Um dieses Ziel zu erreichen, muss der Roboter laufend Vorhersagen über die Position der Orientierungspunkte machen. Die Konfidenzellipsen ergeben sich dabei aus der Annahme von weißem Rauschen in den Bewegungs- und Beobachtungsmodellen.

Unter der Annahme, dass das Rauschen in der UWB-Positionsmessung weißes Rauschen ist, zeigt die Konfidenzellipse die Unsicherheit der Modellvorhersage. Wenn sich das Modell seiner Vorhersage sicher ist, ist die Konfidenzellipse klein. Mit größerer Unsicherheit nimmt auch die Fläche der Konfidenzellipse zu.

Die Modellunsicherheit stammt aus mehreren Hauptquellen, wie z. B. der Unsicherheit, die durch das Explorieren der Agenten entsteht, oder der Unsicherheit, die durch die Unkenntnis des Ground Truth verursacht wird, und der Asynchronität zwischen den einzelnen Agenten, da ein asynchroner Algorithmus zum Trainieren des RL-Modells eingesetzt wird. [27]

Wenn zu einem Zeitpunkt  $t$  eine neue Positionsmessung erfolgt, wird das Modell nicht nur einmal trainiert, sondern  $k$ -mal, sodass auf Grundlage der Modellparameter zum Zeitpunkt  $t - 1$  und der neuen Positionsbeobachtung  $k$ -mögliche Vorhersagen für die Korrektur der Position bestimmt werden. Die einzelnen Vorhersagen werden in einem Vorhersagevektor  $P$  gesammelt. [27]

Zur Messung der Modellunsicherheit wird eine Metrik eingeführt, die auf der Kovarianzmatrix des Vorhersagevektors beruht. Die Kovarianzmatrix  $\Sigma_P$  ist eine zweidimensionale Matrix, die den Unsicherheitsgrad der auf den  $x$ - und  $y$ -Koordinaten angewandten Korrekturmaßnahmen angibt.

Es seien  $a$  und  $b$  die Eigenwerte der Kovarianzmatrix  $\Sigma_P$ . Wenn  $a$  und  $b$  hinreichend von null abweichen, wird das Produkt verwendet, um die Fläche der Konfidenzellipse zu berechnen. Liegt einer der Eigenwerte näher an null, wird die Summe beider Eigenwerte als Ersatzmaß für die Konfidenzellipse eingesetzt. Als Schwellenwert ist  $10^{-7}$  festgelegt. [27]

Algorithmus 1 zeigt die Berechnung des Konfidenzmaßes  $Z$ . Je größer das Konfidenzmaß ist, desto geringer ist die Unsicherheit der Vorhersagen des Agenten. Durch die Maximierung des Konfidenzmaßes wird die Kovarianz der Vorhersagen minimiert. [27]

**Algorithmus 1** Berechnung des Konfidenzmaßes

---

```
while  $i < k$  do                                ▷ Vorhersagevektor  $P$  füllen
  Modell trainieren
  Vorhersage  $p$  berechnen
  Vorhersage  $p$  zu  $P$  hinzufügen
end while

Kovarianzmatrix  $\Sigma_P$  von  $P$  berechnen
Eigenwerte  $a$  und  $b$  von  $\Sigma_P$  berechnen

if  $a \geq 10^{-7}$  &  $b \geq 10^{-7}$  then           ▷ Fläche der Konfidenzellipse bestimmen
   $S_{conf} \leftarrow \pi \cdot a \cdot b$ 
else
   $S_{conf} \leftarrow a + b$ 
end if

return  $Z \leftarrow -S_{conf}$                        ▷ Konfidenzmaß zurückgeben
```

---

Quelle: In Anlehnung an [27]

Das zuvor bestimmte Vertrauensmaß wird verwendet, um die Belohnungsfunktion  $R_t$  zu bilden. Dazu wird das Vertrauensmaß gleichmäßig über die Anzahl  $k$  der ausgeführten Vorhersagen diskontiert, mittelwertbefreit und standardisiert.

Das Ziel des Modells ist es, eine Strategie zu erlernen, die die Belohnungen bei jeder Vorhersage maximiert.

### 5.2.4 Reinforcement Learning Algorithmus

In der Literatur zum Reinforcement Learning sind verschiedene Ansätze, wie z. B. Q-Learning oder Policy-Gradient Methoden bekannt, um die optimale Strategie zu finden. In den letzten Jahren werden besonders auf Deep Learning basierende Algorithmen wie z. B. Deep Q-Learning (DQN), Double Deep Q-Learning (Double DQN) oder Actor-Critic Verfahren eingesetzt, um auch hochdimensionale und kontinuierliche Zustands- und Aktionsräume abstrahieren zu können.

Für das vorgestellte RL-Modell wird die asynchrone Trainingsarchitektur Asynchronous Advantage Actor Critic (A3C) verwendet. A3C ist ein Policy-Gradient Algorithmus, der sowohl eine Strategie als auch eine Wertefunktion erlernt. In einer asynchronen Trainingsarchitektur werden mehrere Agenten parallel in verschiedenen Threads betrieben. Jeder Agent hält eine eigene unabhängige Kopie der Trainingsumgebung und trainiert ein separates Modell. Durch das asynchrone Trainingsprotokoll wird die im DQN notwendige Aktualisierung eines Target-Networks ersetzt. Die Vorteile von A3C sind kürzere Trainingszeiten mit stabileren Ergebnissen gegenüber mehr traditionellen RL-Algorithmen. [17, 27]

Der Agent quantifiziert den Advantage als ein Maß, das eine Aussage über die Qualität der Aktionen liefert. Ähnlich wie beim DQN wird der Advantage wie folgt berechnet:

$$A(s, a) = Q(s, a) - V(s) \quad (5.1)$$

In A3C kann der Advantage angenähert werden, indem  $Q(s, a)$  durch  $R_t$  ersetzt wird. Folglich gilt für den Advantage:

$$A(s, a) = R_t - V(s) \quad (5.2)$$

Das Ziel von A3C ist es, den gesamten Verlust (engl. loss) zu minimieren, der sich aus dem Strategieverlust (engl. policy loss), dem Verlust der Wertefunktion (engl. value loss) und dem Entropieverlust (engl. entropy loss) zusammensetzt. Es sei  $J(\pi_\theta)$  definiert als die Gesamtbelohnung, die ein Agent mit der Strategie  $\pi$  erhält. Aus dem Policy-Gradient Theorem folgt:

$$\nabla_\theta J(\pi_\theta) = \mathbb{E}[A(s, a) \nabla_\theta \log \pi_\theta(a | s)], \quad (5.3)$$

wobei  $\nabla_\theta J(\pi_\theta)$  der Gradient der Belohnungsfunktion  $J(\pi_\theta)$  ist und  $\mathbb{E}[A(s, a) \nabla_\theta \log \pi_\theta(a | s)]$  der Erwartungswert, der den approximierten Policy-Gradienten angibt. Wenn  $A(s, a)$  als konstant betrachtet wird, ist der Strategieverlust  $L_\pi$  definiert als:

$$L_\pi = -\frac{1}{n} \sum_{i=1}^n \log \pi_\theta(a_i | s_i) A(s_i, a_i) \quad (5.4)$$

Mit A3C wird zudem eine Wertfunktion gelernt. Da die tatsächliche Wertfunktion der Bellman-Gleichung entspricht, ist der Fehler  $e$  zwischen der in  $n$ -Schritten approximierten Wertfunktion und der wahren Wertfunktion:

$$e = r_0 + \gamma r_1 + \gamma^2 r_2 + \dots + \gamma^n V(s_n) - V(s_0) \quad (5.5)$$

Der Verlust der Wertfunktion  $L_V$  ist definiert als:

$$L_V = \frac{1}{n} \sum_{i=1}^n e_i^2 \quad (5.6)$$

Der Entropieverlust wird als Regularisierungsterm verwendet, der wie folgt definiert ist

$$L_{ent} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \pi_\theta(s_i)_j \log \pi_\theta(s_i)_j \quad (5.7)$$

Der gesamte Verlust  $L$  setzt sich schließlich folgendermaßen zusammen:

$$L = L_\pi + L_V + \lambda L_{ent} \quad (5.8)$$

### 5.2.5 Modellarchitektur

Das Modell nimmt in jedem Schritt eine Matrix von  $x$ - und  $y$ -Koordinaten als Eingabe entgegen. Die Größe der Eingabe wird als  $50 \times 2$ -Matrix festgelegt, die 49 historische Modellvorhersagen und die aktuell gemessene Position enthält.

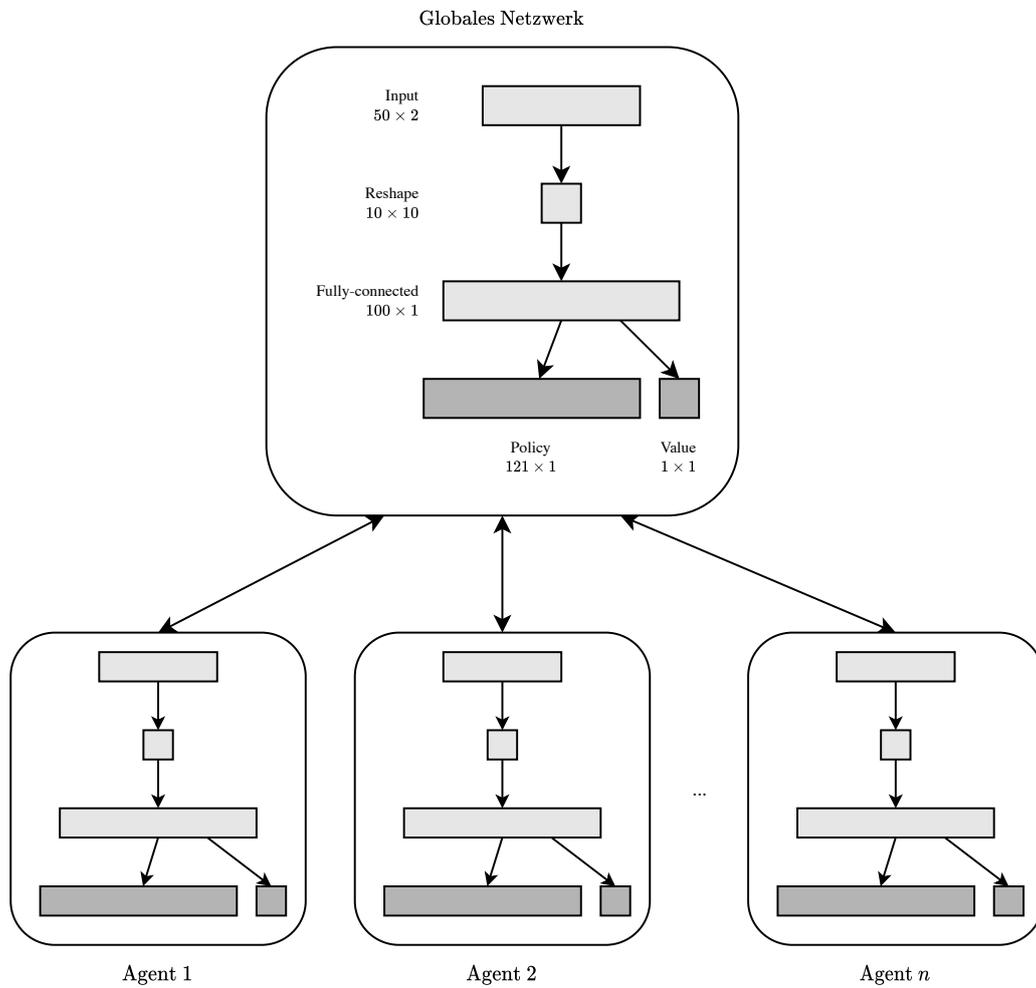


Abbildung 5.5: A3C Modellarchitektur

Quelle: In Anlehnung an [27]

Die Modellarchitektur ist in Abbildung 5.5 dargestellt. Auf den  $50 \times 2$  Eingabe-Layer folgt ein Skalierungs-Layer, der die Form in einen  $10 \times 10$ -Layer umwandelt. Anschließend folgt ein Dense-Layer der Dimension  $100 \times 1$  mit ReLU-Aktivierungsfunktion. Mit den Ausgängen werden zwei Ausgabe-Layer gebildet. Die Policy-Ausgabe ist ein 121-dimensionaler Dense-Layer mit Softmax-Aktivierungsfunktion, während die Value-Ausgabe ein eindimensionaler Dense-Layer mit linearer Aktivierungsfunktion ist.

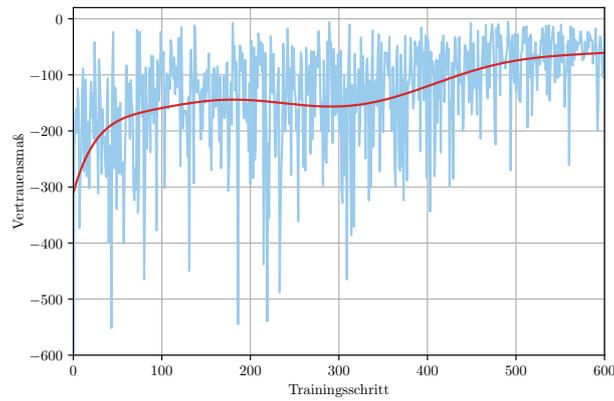
Alle parallelen Agenten haben die gleiche Netzwerkarchitektur wie das globale Netzwerk. Die von den einzelnen Agenten berechneten Gradienten werden gesammelt und verwendet, um das gemeinsame globale Netzwerk zu aktualisieren.

Das RL-Modell wurde mit Python, Tensorflow und Keras implementiert. [5]

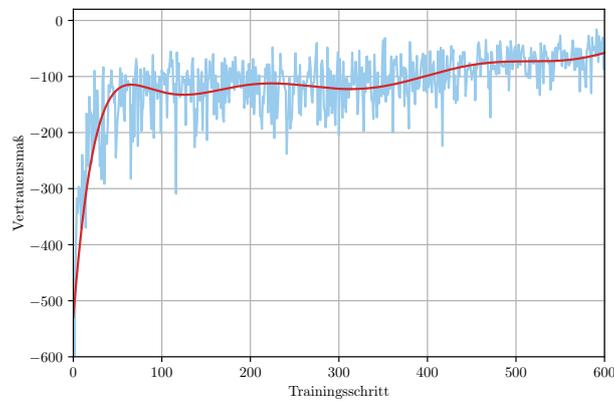
Zum Training des neuronalen Netzes wird der RMSprop-Optimizer verwendet. Die Lernrate wird auf 0,01 und der entsprechende Decay-Parameter auf den Wert 0,99 gesetzt. Die Anzahl der Vorhersagen, die zur Berechnung des Vertrauensmaßes verwendet werden, beträgt 30; d. h.  $k = 30$ . Die Vorhersage für jeden neuen beobachteten Punkt ist der Mittelpunkt  $(\bar{x}, \bar{y})$  der Konfidenzellipse.

Da das Ziel des RL-Modells darin besteht, die Belohnung (bzw. das Vertrauensmaß) zu maximieren, wird die Leistung des Modells unter anderem anhand dieser Metrik bewertet. Um die Entwicklung des Vertrauensmaßes während des Trainings zu beobachten, wird für jeden Trainingsschritt der entsprechende Wert protokolliert.

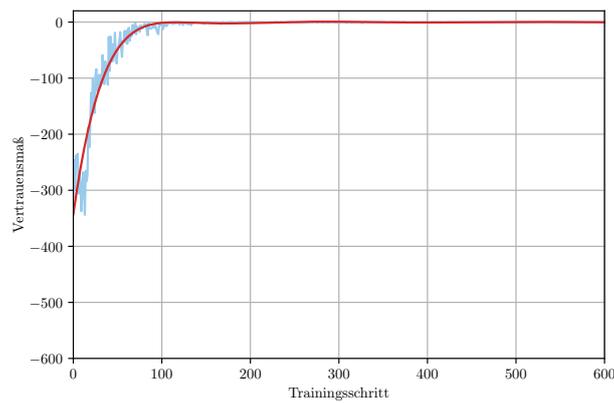
Die Anzahl der für das Training verwendeten Threads kann die Ergebnisse bezüglich der Konvergenzrate und der optimalen Strategie beeinflussen. Um die Auswirkungen zu vergleichen, wurde der Trainingsalgorithmus mit einem einzelnen Thread, mit zwei Threads und mit vier Threads ausgeführt. Die Ergebnisse sind in Abbildung 5.6 dargestellt. Es ist deutlich erkennbar, dass im Training mit vier Threads das beste Ergebnis erzielt wurde. Das Vertrauensmaß konvergiert schnell gegen einen hohen Wert bei geringer Streuung. Durch die parallele Trainingsarchitektur wird die Diversität erhöht, sodass in den vier Trainingsumgebungen vier unabhängige Szenarien gelernt werden können. Bei einem Training mit nur einem Thread zeigen die Experimente, dass das Modell nur langsam konvergiert. Auch mit zwei parallelen Threads wird nach 600 Trainingsschritten noch keine vollständige Konvergenz erreicht. Zudem ist die Standardabweichung des Vertrauensmaßes in den Trainingsdurchgängen mit einem und zwei Threads wesentlich höher. Eine größere Standardabweichung deutet auf eine größere Unsicherheit bezüglich der erlernten Strategie hin.



(a) 1 Thread; mittleres Vertrauensmaß =  $-118,1$ ;  
mittlere Standardabweichung =  $72,8$



(b) 2 Threads; mittleres Vertrauensmaß =  $-101,6$ ;  
mittlere Standardabweichung =  $34,7$



(c) 4 Threads; mittleres Vertrauensmaß =  $-12,2$ ;  
mittlere Standardabweichung =  $3,3$

Abbildung 5.6: Sensitivitätsanalyse der Anzahl der im Training verwendeten Threads

Quelle: Eigene Darstellung

Für den produktiven Einsatz des RL-Agenten wurden auf Grundlage dieser Beobachtungen vier Threads für die A3C Trainingsarchitektur ausgewählt. In Abbildung 5.6c ist zu erkennen, dass mit dieser Wahl bereits nach ungefähr 100 Trainingsschritten die vollständige Konvergenz erreicht ist.

Auch wenn der RL-Agent im Training erlernt das Konfidenzmaß zu maximieren, müssen die resultierenden Positionskorrekturen des Agenten nicht unbedingt mit dem exakten Verlauf der Ground Truth Positionen übereinstimmen. Zhang und Masoud [27] beschreiben dieses Problem ebenfalls in ihrer Arbeit. Als Lösung stellen die Autoren ein Map-Matching Algorithmus vor, der als Regularisierungsparameter in die Belohnungsfunktion aufgenommen wird. Das Map-Matching setzt jedoch voraus, dass eine entsprechende Karte existiert und kann daher nicht auf den Anwendungsfall des UWB-Lokalisationssystems übertragen werden.

### 5.3 Kalman-Filter

Der Kalman-Filter ist ein Verfahren zur Abschätzung von Systemzuständen. Dabei arbeitet der Filter in zwei Schritten. Im ersten Schritt, der Vorhersage, wird der neue Systemzustand anhand früherer Informationen geschätzt, ohne dass eine neue Messung vorliegt. Im zweiten Schritt, der Korrektur, wird diese Schätzung zusammen mit möglicherweise rauschbehafteten Messdaten verwendet, um eine optimale Vorhersage für den aktuellen Systemzustand zu treffen. [26]

Für den Einsatz im UWB-Lokalisationssystem soll der Kalman-Filter den Verlauf der gesuchten Position abschätzen. Der Systemzustand  $\mathbf{x}$  setzt sich daher aus einer  $x$ - und einer  $y$ -Koordinate sowie deren zwei zeitlichen Ableitungen zusammen.

$$\mathbf{x} = \begin{pmatrix} x & \dot{x} & \ddot{x} & y & \dot{y} & \ddot{y} \end{pmatrix}^\top$$

Die Messung  $\mathbf{z}$  enthält die vom Lokalisationssystem erfassten  $x$ - und  $y$ -Koordinaten.

$$\mathbf{z} = \begin{pmatrix} x & y \end{pmatrix}^\top$$

Da das UWB-Lokalisationssystem die  $x$ - und  $y$ -Positionen direkt bestimmt, ist die Modellierung der Beobachtungsmatrix  $\mathbf{H}$  trivial. Die  $x$ - und  $y$ -Koordinaten werden direkt aus  $\mathbf{z}$  entgegengenommen. Die Ableitungen sollen hingegen indirekt durch den Filter bestimmt werden.

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Das Modell zur Vorhersage des Systemzustands folgt der klassischen Mechanik. Nach den Newtonschen Grundgesetzen der Bewegung kann eine Vorhersage der Position im nächsten Zeitschritt berechnet werden. Für die Berechnung der Position, Geschwindigkeit und Beschleunigung in  $x$ - und  $y$ -Richtung ergeben sich folgende Differentialgleichungen:

$$x(t_n) = x(t_{n-1}) + \Delta t \dot{x}(t_{n-1}) + \frac{\Delta t^2}{2} \ddot{x}(t_{n-1}) + \frac{\Delta t^3}{6} J_x \quad (5.9a)$$

$$\dot{x}(t_n) = \dot{x}(t_{n-1}) + \Delta t \ddot{x}(t_{n-1}) + \frac{\Delta t^2}{2} J_x \quad (5.9b)$$

$$\ddot{x}(t_n) = \ddot{x}(t_{n-1}) + \Delta t J_x \quad (5.9c)$$

$$y(t_n) = y(t_{n-1}) + \Delta t \dot{y}(t_{n-1}) + \frac{\Delta t^2}{2} \ddot{y}(t_{n-1}) + \frac{\Delta t^3}{6} J_y \quad (5.9d)$$

$$\dot{y}(t_n) = \dot{y}(t_{n-1}) + \Delta t \ddot{y}(t_{n-1}) + \frac{\Delta t^2}{2} J_y \quad (5.9e)$$

$$\ddot{y}(t_n) = \ddot{y}(t_{n-1}) + \Delta t J_y \quad (5.9f)$$

Dabei ist  $t_n$  der  $n$ -te Zeitschritt. Zudem werden die Terme höherer Ordnung (einschließlich des Rucks)  $J_x$  und  $J_y$  als normalverteilt angenommen.

Mit diesem Modell kann aus dem aktuellen Zustand ein Folgezustand vorhergesagt werden. Aufgrund der Dynamik in den Termen höherer Ordnung und der mit dem Zeitschritt verbundene Fehler ist das Modell auf eine Vorhersage weniger Zeitschritte begrenzt. [25]

In Matrixschreibweise ergibt sich für das zuvor eingeführte Newtonsche Modell die Propagationsmatrix

$$\mathbf{F} = \begin{pmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} & 0 & 0 & 0 \\ 0 & 1 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 0 & 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Zu beachten ist, dass nur zwei Zeitableitungen verwendet werden und die Beschleunigung als konstant betrachtet wird.

Mit der Matrix  $\mathbf{F}$  wird der vorherige Zustand auf den Systemzustand zum aktuellen Zeitpunkt propagiert. Es gilt  $\mathbf{x}(t_n) \approx \mathbf{F} \cdot \mathbf{x}(t_{n-1})$ . [25]

Für die Modellierung des Prozessrauschens  $\mathbf{Q}$  wird der Einfachheit halber angenommen, dass das Rauschen für jede Zeitperiode konstant ist. Durch die Annahme dieses Piecewise White Noise Modells wird es ermöglicht, eine Varianz zu verwenden, um anzugeben, wie sehr sich das System zwischen den Schritten ändert. [15]

Als Varianz eignen sich etwa die normalverteilten Werte der dritten Zeitableitung der  $x$ - und  $y$ -Positionen. [25]

$$\mathbf{Q} = \sigma_J^2 \begin{pmatrix} \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} & \frac{\Delta t^2}{2} & 0 & 0 & 0 \\ \frac{\Delta t^3}{2} & \Delta t^2 & \Delta t & 0 & 0 & 0 \\ \frac{\Delta t^2}{2} & \Delta t & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} & \frac{\Delta t^2}{2} \\ 0 & 0 & 0 & \frac{\Delta t^3}{2} & \Delta t^2 & \Delta t \\ 0 & 0 & 0 & \frac{\Delta t^2}{2} & \Delta t & 1 \end{pmatrix}$$

Das Messrauschen in der Positionsbestimmung des UWB-Lokalisationssystems wird auf die Herstellerangabe aus dem Datenblatt des UWB-Transceivers von  $\pm 10$  cm festgelegt. Die Fehlerkovarianzmatrix  $\mathbf{R}$  ist somit definiert als

$$\mathbf{R} = \begin{pmatrix} 10^2 & 0 \\ 0 & 10^2 \end{pmatrix}.$$

Der Vorhersagefehler für  $\hat{\mathbf{x}}$  wird im gegebenen Anwendungsfall mit einem Startwert von 10 belegt. Die entsprechende Kovarianzmatrix  $\mathbf{P}$  wird durch Multiplikation mit der Einheitsmatrix erhalten.

$$\mathbf{P} = 10 \cdot \mathbf{I}$$

Nach der passenden Initialisierung der Variablen kann der Kalman-Filter in einem iterativen und rekursiven Prozess abwechselnd den Systemzustand vorausberechnen und mit den jeweils aktuellen Messwerten korrigieren.

Während des ersten Schrittes wird mithilfe des modellierten Systemverhaltens und des vorhergehenden Zustands eine Vorhersage für den aktuellen Zeitschritt gemacht. Die vorausgesagte Positionsangabe bedingt vom Zeitschritt  $t_{n-1}$  ist in der Zustandsschätzung  $\hat{\mathbf{x}}(t_n | t_{n-1})$  verfügbar. [15]

$$\hat{\mathbf{x}}(t_n | t_{n-1}) = \mathbf{F} \cdot \hat{\mathbf{x}}(t_{n-1}) \quad (5.10a)$$

$$\mathbf{P}(t_n | t_{n-1}) = \mathbf{F} \cdot \mathbf{P}(t_{n-1}) \cdot \mathbf{F}^\top + \mathbf{Q} \quad (5.10b)$$

Die Vorhersagen aus dem ersten Schritt werden anschließend mit den tatsächlichen Messwerten korrigiert. Durch  $\hat{\mathbf{x}}(t_n)$  wird der aktuell geschätzte Systemzustand, somit auch die gesuchte Position, angegeben. [15]

$$\mathbf{K}(t_n) = \mathbf{P}(t_n | t_{n-1}) \cdot \mathbf{H}^\top \cdot (\mathbf{H} \cdot \mathbf{P}(t_n | t_{n-1}) \cdot \mathbf{H}^\top + \mathbf{R})^{-1} \quad (5.11a)$$

$$\hat{\mathbf{x}}(t_n) = \hat{\mathbf{x}}(t_n | t_{n-1}) + \mathbf{K}(t_n) \cdot (\mathbf{z}(t_n) - \mathbf{H} \cdot \hat{\mathbf{x}}(t_n | t_{n-1})) \quad (5.11b)$$

$$\mathbf{P}(t_n) = (\mathbf{I} - \mathbf{K}(t_n) \cdot \mathbf{H}) \cdot \mathbf{P}(t_n | t_{n-1}) \quad (5.11c)$$

Einer der Hauptvorteile des Kalman-Filters ist die Berechnungseffizienz bei der Ermittlung einer optimalen Lösung. Allerdings sind die Voraussetzungen, unter denen ein Kalman-Filter die optimale Lösung bestimmen kann, in der Praxis nicht immer erfüllt, da die Modellierung des Systems auf lineare Zustandsräume beschränkt ist. Um einen

Kalman-Filter auch in nichtlinearen Systemen anwenden zu können, müssen lokale Linearisierungsverfahren verwendet werden. Der Erweiterte Kalman-Filter (EKF) oder der Unscented Kalman-Filter (UKF) verwenden verschiedene Linearisierungstechniken zur Lösung von nichtlinearen Zustandsgleichungen. [20]

Diese Erweiterungen des klassischen Kalman-Filters werden in dieser Arbeit jedoch nicht weiter behandelt.

### 5.4 Particle-Filter

Particle-Filter, auch bekannt als Sequenzielle Monte-Carlo-Methode (SMC), stellen die posteriore Wahrscheinlichkeit eines Systemzustands zu einem bestimmten Zeitpunkt durch eine Reihe von Stichproben, den Partikeln, und zugehörige Gewichte, die Wichtigkeitsfaktoren, dar. [20]

Der Hauptvorteil eines Particle-Filters ist die Fähigkeit, jede beliebige Wahrscheinlichkeitsverteilung abzubilden. Im Gegensatz zum Kalman-Filter kann der Particle-Filter auch Zustandsschätzungen für nichtlineare und nicht-gaußsche Zusammenhänge bereitstellen. Dabei arbeitet der Particle-Filter nach einer Monte Carlo (MC) Methode. Es werden ausreichend viele Partikel erzeugt, um eine repräsentative Stichprobe des Problems zu erhalten. [15]

Ein Particle-Filter arbeitet allgemein nach folgendem Ablauf:

1. Partikel erzeugen.

Die gesuchte Position ist anfangs unbekannt. Es wird eine Menge zufälliger Partikel verteilt über die gesamte Umgebung erzeugt. Jeder Partikel wird zunächst mit den gleichen Gewichten initialisiert.

2. Vorhersage.

Die Partikel werden passend zu den Änderungen des realen System angepasst. Dabei wird den Partikeln ein Rauschen hinzugefügt, um die Unsicherheit des realen Systems zu modellieren.

3. Aktualisieren.

Die aktuelle Messung wird verwendet, um die Wahrscheinlichkeitsverteilung zu aktualisieren. Partikel, die gut mit der Messung übereinstimmen, werden höher gewichtet als Partikel, die weniger gut mit der Messung übereinstimmen.

### 4. Resampling.

Beim Resampling werden Partikel verworfen, die sehr unwahrscheinlich sind und durch Partikel mit höherer Wahrscheinlichkeit ersetzt.

### 5. Zustandsschätzung berechnen.

Durch die Berechnung des gewichteten Mittelwerts und der Kovarianz der Partikelmenge wird die aktuelle Zustandsschätzung ermittelt.

Ein Particle-Filter wird als ein weiterer Algorithmus zur Verbesserung der Positionierungsgenauigkeit erprobt. Die verwendete Implementation arbeitet mit einer Gesamtzahl von 3000 Partikeln.

## 5.5 Vergleich

Zur Auswertung wurden im CSTI gemeinsam mit dem AR-Trackingsystem zwei Trajektorien aufgenommen. Die Abbildungen 5.7 und 5.8 enthalten die Messpunkte des AR-Trackingsystems (Ground Truth) und die Rohdaten des UWB-Lokalisationsystems. Abbildung 5.7 zeigt den Ausschnitt einer möglichst geraden Strecke, während in Abbildung 5.8 ein kreisförmiger Bewegungspfad ausgewählt wurde. Die drei zuvor vorgestellten Verfahren zur Erhöhung der Genauigkeit werden auf die Rohdaten des UWB-Lokalisationsystems angewendet und die resultierenden korrigierten Positionen mit den Ground Truth Daten verglichen.

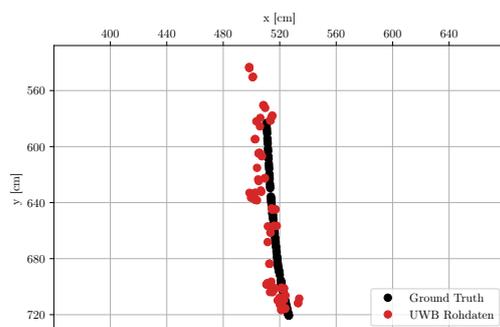


Abbildung 5.7: Ausschnitt einer geradlinigen Trajektorie

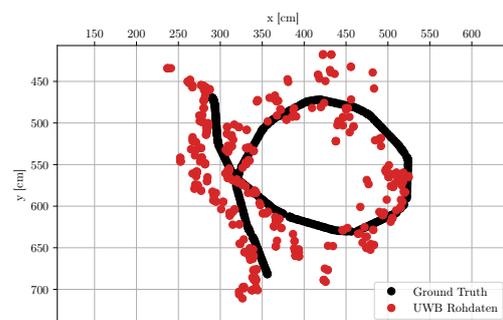


Abbildung 5.8: Ausschnitt einer kreisförmigen Trajektorie

Quelle: Eigene Darstellung

Zur Bewertung der Leistungsfähigkeit der drei Algorithmen wird der Fehler für jede einzelne Positionsmessung und der akkumulierte Fehler über die gesamte Messreihe betrachtet. Für jede Positionsmessung  $i$  ist  $(g_{xi}, g_{yi})$  als Ground Truth und  $(l_{xi}, l_{yi})$  als der vorhergesagte Punkt des Korrekturverfahrens definiert. Der Fehler  $E_i$  gibt die Abweichung in jedem Zeitschritt  $i$  an und wird wie folgt berechnet:

$$E_i = \sqrt{(g_{xi} - l_{xi})^2 + (g_{yi} - l_{yi})^2} \quad (5.12)$$

Der akkumulierte Fehler  $E_{all}$  ist die Summe aller vorherigen Fehler:

$$E_{all} = \sum_{i=0}^C E_i \quad (5.13)$$

Dabei ist  $C$  die Anzahl der Messungen. Der Vergleich dieser Fehlergrößen ermöglicht die Beurteilung der Leistungsfähigkeit der einzelnen Algorithmen. In den Abbildungen 5.9 bis 5.12 sind die Fehlermaße für beide Trajektorien und die einzelnen Korrekturverfahren graphisch dargestellt.

Beim Vergleich des geradlinigen und kreisförmigen Verlaufs untereinander fällt zunächst auf, dass die Fehler der kreisförmigen Trajektorie größer sind als bei der geraden Strecke. Bereits die Rohdaten aus dem UWB-Lokalisationssystem weichen hier stärker vom Ground Truth ab. Die Ursache für dieses Phänomen konnte nicht abschließend geklärt werden. Eine mögliche Begründung ist das verstärkte Auftreten von Effekten wie beispielsweise der Mehrwegeausbreitung, wodurch das Messrauschen und die Ungenauigkeit zunehmen.

Es lässt sich weiterhin feststellen, dass alle der getesteten Algorithmen geeignet sind, um die Positionsmessungen des UWB-Lokalisationssystems zu verbessern. Einerseits der durchschnittliche Fehler pro Zeitschritt und andererseits der akkumulierte Fehler befinden sich für beide Trajektorien deutlich unterhalb des Fehlers der Rohdaten.

Nach der zuvor definierten Fehlermetrik erzielt der Particle-Filter sowohl für die geradlinige als auch für die kreisförmige Trajektorie die besten Ergebnisse. Der durchschnittliche Fehler des Particle-Filters pro Zeitschritt beträgt 9,06 cm bzw. 28,87 cm.

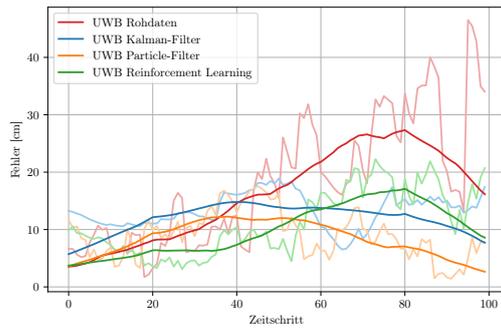


Abbildung 5.9: Fehler der geradlinigen Trajektorie

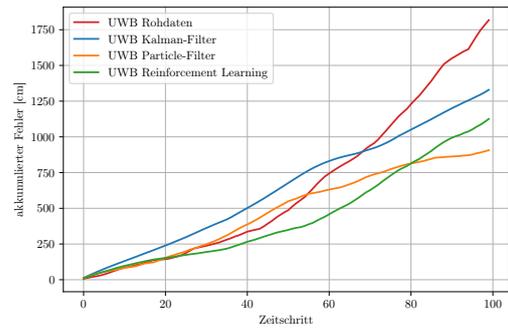


Abbildung 5.10: Akkumulierter Fehler der geradlinigen Trajektorie

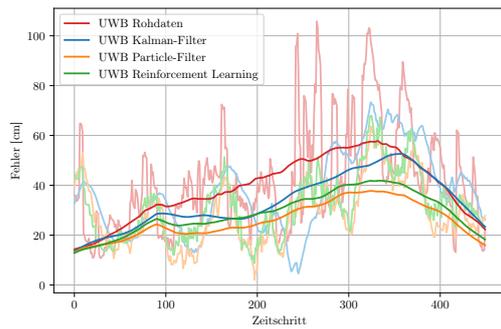


Abbildung 5.11: Fehler der kreisförmigen Trajektorie

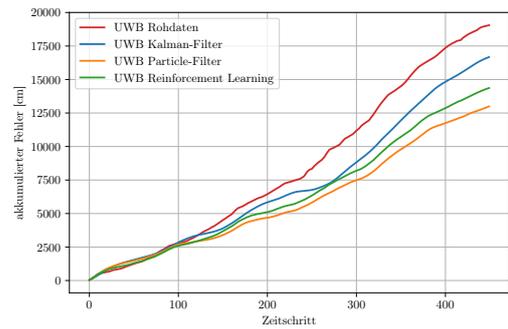


Abbildung 5.12: Akkumulierter Fehler der kreisförmigen Trajektorie

Quelle: Eigene Darstellung

Mit diesem Ergebnis kann zumindest für die geradlinige Trajektorie die Angabe des Herstellers *Decawave* in Bezug auf eine Genauigkeit des UWB-Transceivers von  $\pm 10$  cm in Lokalisierungsanwendungen bestätigt und reproduziert werden. Bezugnehmend auf die in Kapitel 2 vorgestellten Genauigkeitsanforderungen an unterschiedliche Lokalisierungsanwendungen lässt sich auch die in der kreisförmigen Bewegung erreichte Genauigkeit positiv bewerten. Eine Genauigkeit von  $\pm 30$  cm ist für viele Anwendungsfälle ausreichend.

Die Leistung des RL-Modells befindet sich nur sehr knapp hinter den Ergebnissen des Particle-Filters. Im Experiment mit der geradlinigen Strecke beträgt der Fehler 11,26 cm und für die kreisförmige Bahn 31,92 cm.

Die Leistung des Kalman-Filters ist im Vergleich zu den vorherigen zwei Ansätzen etwas schlechter. Dennoch konnten die Positionsbestimmungen in beiden Experimenten verbessert werden und der durchschnittliche Fehler für den geraden Verlauf auf 13,29 cm und für die ringförmige Trajektorie auf 37,34 cm reduziert werden.

Diese Ergebnisse lassen sich auch bei Betrachtung der Abbildungen 5.13 bis 5.18, die die Verläufe der korrigierten Trajektorien darstellen, wiedererkennen.

Auffällig bei der Analyse der durch den Kalman-Filter korrigierten Positionsmessungen (vgl. Abbildung 5.15 und 5.16) ist, dass die Trajektorien deutlich geglättet sind, während bei den probabilistischen Verfahren ein Rauschen zu erkennen ist.

Durch den geglätteten Verlauf ist einerseits die Varianz der korrigierten Positionen sehr gering. Andererseits liegt die korrigierte Trajektorie stellenweise nicht über dem Verlauf der tatsächlichen Ground Truth Daten. Diese Abweichung wird insbesondere in Abbildung 5.16 deutlich. In der durch den Kalman-Filter korrigierten kreisförmigen Trajektorie ist eine deutliche Abweichung erkennbar. Der kreisförmige Verlauf ist nicht optimal durch das im Kalman-Filter verwendete Bewegungsmodell vorhersagbar. Wie eingangs vermutet, können in diesem Fall das RL-Modell und der Particle-Filter (vgl. Abbildung 5.14 und 5.18) mit ihrer Fähigkeit auch nichtlineare und nicht-gaußsche Zusammenhänge abzuschätzen, eine deutlich bessere Vorhersage treffen.

Wird zusätzlich zur Korrekturleistung der Rechenaufwand mit in die Bewertung einbezogen, profitiert der Kalman-Filter durch die recht einfache Berechenbarkeit in Form von wenigen Matrixoperationen. Diese Berechnungen könnten problemlos auch auf einem Mikrocontroller mit Gleitkommaeinheit (engl. floating point unit) implementiert werden. Für das Training des RL-Modells mit der asynchronen Trainingsarchitektur ist hingegen deutlich mehr Rechenleistung notwendig. Zudem profitiert der parallele A3C Trainingsalgorithmus von einem Mehrkernprozessor. Auch der Particle-Filter, bei dem tausende

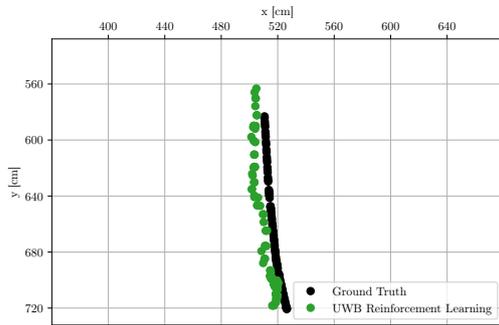


Abbildung 5.13: Geradlinige Trajektorie korrigiert mit dem RL-Modell

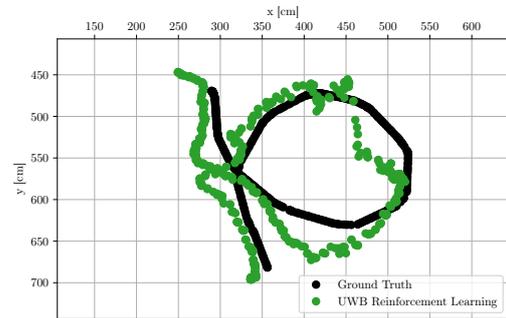


Abbildung 5.14: Kreisförmige Trajektorie korrigiert mit dem RL-Modell

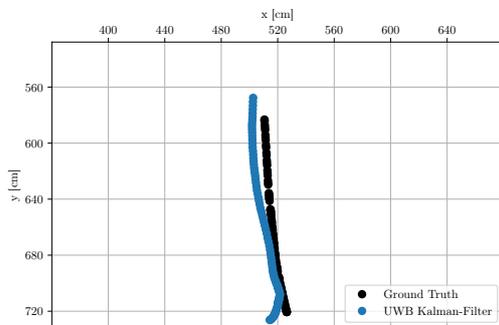


Abbildung 5.15: Geradlinige Trajektorie korrigiert mit dem Kalman-Filter

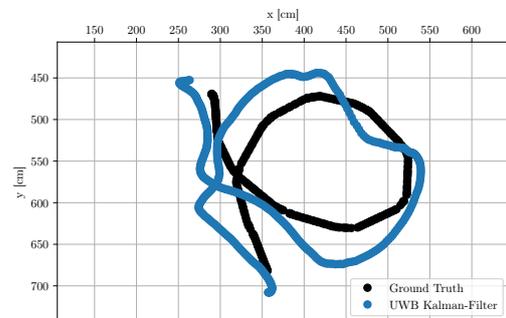


Abbildung 5.16: Kreisförmige Trajektorie korrigiert mit dem Kalman-Filter

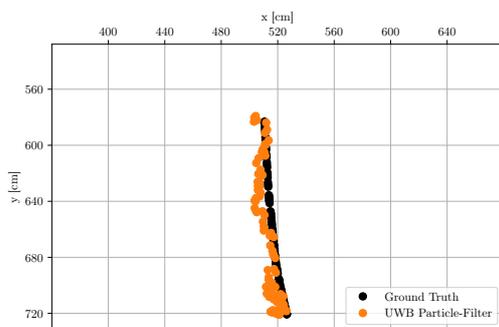


Abbildung 5.17: Geradlinige Trajektorie korrigiert mit dem Particle-Filter

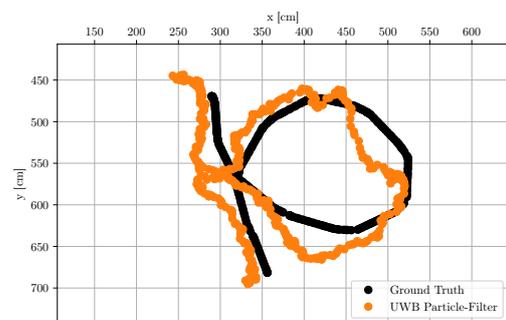


Abbildung 5.18: Kreisförmige Trajektorie korrigiert mit dem Particle-Filter

Quelle: Eigene Darstellung

gewichtete Systemzustände ausgewertet und aktualisiert werden müssen, ist rechenintensiv. Für einen Particle-Filter steigt die Rechenkomplexität exponentiell zur Anzahl der Elemente im Zustandsvektor [20].

In eingebetteten oder tragbaren Systemen, die zur Zielgruppe des UWB-Lokalisationsystems zählen, stehen gewöhnlich nur in begrenztem Maße Rechenleistung und Energie zur Verfügung. Der Einsatz von komplexen Algorithmen zur Positionierung sollte daher sorgfältig überprüft werden.

Die Experimente zeigen weiterhin, dass neben den etablierten Filter- und Korrekturalgorithmen auch Reinforcement Learning zur Verbesserung der Lokalisierungsgenauigkeit eingesetzt werden kann. Das vorgestellte RL-Modell konvergiert schnell und erreicht Ergebnisse, die gleichwertig oder zum Teil besser als die traditionellen Methoden sind.

## 6 Fazit

In dieser Arbeit wurde ein Lokalisationssystem basierend auf der Ultra-wideband Technologie vorgestellt. Das System umfasst eine individuell entwickelte Hardwareplattform, die Implementation eines DS-TWR Protokolls zur drahtlosen Entfernungsmessung und die Realisierung einer zentralen Lokalisierungseinheit zur Berechnung und Korrektur der gesuchten Position.

Die Anforderungen an Indoor- und Outdoor-Positionierungssystemen sind sehr divers. In Innenräumen gestaltet sich eine präzise Positionsbestimmung oft schwieriger, da aufgrund von Hindernissen und Störquellen die Signalqualität von insbesondere schmalbandigen Drahtlossystemen stark abnehmen kann. UWB-Signale zeichnen sich hingegen durch eine besonders hohe Bandbreite aus. Mit dieser Eigenschaft eignet sich Ultra-wideband besonders für Entfernungs- und Positionsmessungen.

Es existieren verschiedene Möglichkeiten und Verfahren zur Bestimmung der Position mithilfe von Drahtlossystemen. Besonders herausgearbeitet wurde der ToA Signalparameter und die Trilateration als geometrisches Verfahren zur Berechnung der Position. Das vorgestellte UWB-Lokalisationssystem nutzt diese zwei Verfahren für eine zuverlässige und robuste Positionsbestimmung. Die Implementation des Lokalisationssystems berücksichtigt dabei internationale Standards wie beispielsweise IEEE 802.15.4.

Zur Korrektur der Fehler und Messunsicherheiten werden drei Algorithmen vorgestellt. Der Kalman-Filter ist eine gängige Methode zur Erhöhung der Genauigkeit in Positionierungssystemen. Zudem wird mit dem Particle-Filter eine etablierte Lösung aus der Robotik betrachtet. Als modernes Verfahren wird zudem ein auf Reinforcement Learning basierendes Verfahren erprobt, um die Genauigkeiten der Positionsbestimmungen zu erhöhen. Das RL-Modell verwendet die letzten Vorhersagen und die vom Lokalisationssystem gemessene Position als Eingabe, sodass auf dieser Grundlage eine optimale Strategie zur Positionskorrektur erlernt werden kann. Für den Belohnungsmechanismus wird ein Vertrauensmaß eingesetzt, das aus mehreren Modellvorhersagen errechnet wird.

Durch den Einsatz der parallelen A3C Trainingsarchitektur wird eine schnelle Konvergenz des RL-Modells erreicht.

Im Vergleich der drei Korrekturansätze hat sich gezeigt, dass alle getesteten Algorithmen geeignet sind, die Positionsmessungen zu verbessern. Besonders gute Ergebnisse werden unter Verwendung des Particle-Filters und des RL-Modells erreicht. Die Leistung des Kalman-Filters nimmt bei nichtlinearen Bewegungen ab, hat jedoch größere Vorteile durch eine geringere Rechenkomplexität.

Mit der im Experiment erreichten Genauigkeit von  $\pm 10$  cm (resp.  $\pm 30$  cm) ist das vorgestellte UWB-Lokalisationssystem in den meisten Anwendungsszenarien geeignet einsetzbar.

Es bleibt zu erproben, inwiefern das vorgestellte Lokalisationssystem mit einer Genauigkeit von höchstens  $\pm 10$  cm für die Miniaturautonomie geeignet ist. Verglichen zur Größe der Modellfahrzeuge und -umgebung erscheint diese Genauigkeit als zu gering. Auf den Modellbaumaßstab H0 (1:87) umgerechnet, ergibt sich jedoch eine Genauigkeit von  $\pm 8,7$  m, was damit noch im Bereich der Genauigkeit von einfachen GPS-Systemen fällt. [14]

Grundsätzlich hängt die Einsatzmöglichkeit von Ultra-wideband in der Miniaturautonomie von dem spezifischen Anwendungsfall an. Denkbar sind unter anderem auch Anwendungen, in denen keine akkurate Positionsauflösung erforderlich ist.

Insgesamt konnte das Ziel der Arbeit erreicht werden, indem die Grundlage für ein auf Ultra-wideband basierendes Lokalisationssystem entwickelt wurde. Nach einigen individuellen Anpassungen könnte das System bereits testweise in Forschungsprojekten an der HAW Hamburg zum Einsatz kommen.

### 6.1 Ausblick

Die Arbeit zeigt, dass das vorgestellte UWB-Lokalisationssystem in verschiedenen Anwendungsfällen einsetzbar ist. Für eine zuverlässige und dauerhafte Nutzung des Systems sind jedoch noch Verbesserungen notwendig. Abschließend werden einige Vorschläge für nützliche Optimierungen aufgeführt.

Zur weiteren Verbesserung der Lokalisierungsgenauigkeit ist die Sensordatenfusion der UWB-Messungen mit zusätzlichen Sensorwerten denkbar. Dafür bieten sich beispielsweise Inertialmessungen einer IMU oder die Odometriedaten von Fahrplattformen an. Zudem ist aber auch die Einbindung von komplexerer Sensorik wie z. B. Radar, Lidar oder Tiefenkameras denkbar, um ein robustes und zuverlässiges Positionierungssystem umzusetzen.

Eine nützliche Weiterentwicklungsmöglichkeit stellt die Ausweitung der Positionsbestimmung im dreidimensionalen Raum dar. Damit würde das UWB-Lokalisationssystem beispielsweise auch für Drohnen und Multicopter nutzbar werden.

In einer zweiten Hardwarerevision ist geplant, die UWB-Beacons zusätzlich mit einer Ethernet-Schnittstelle auszustatten. Damit wird die Umsetzung von bekannten Algorithmen zur Zeitsynchronisation vereinfacht. Mit einer synchronisierten Uhr über die UWB-Anchors hinweg kann die Distanzmessung mittels des TDoA Signalparameters umgesetzt werden, was zu deutlichen Leistungsverbesserungen führt. Für TDoA Entfernungsmessungen werden wesentlich weniger Nachrichten benötigt, sodass eine energiesparendere Umsetzung, höhere Aktualisierungsraten und mehrere Teilnehmer ermöglicht werden.

Das eingesetzte RL-Modell zur Verbesserung der Lokalisierungsgenauigkeit könnte hinsichtlich der Vorhersage zukünftiger Trajektorien erweitert werden. Mit der aktuellen Umsetzung können nur gemessene Koordinaten korrigiert werden, wodurch das RL-Modell keine Vorhersagen treffen kann, wenn das UWB-Lokalisationssystem kurzzeitig ausfällt.

# Literaturverzeichnis

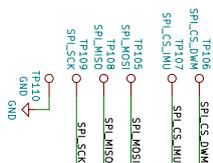
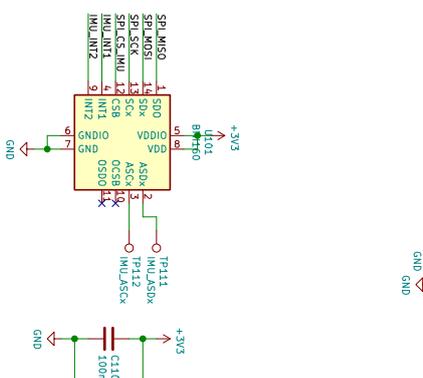
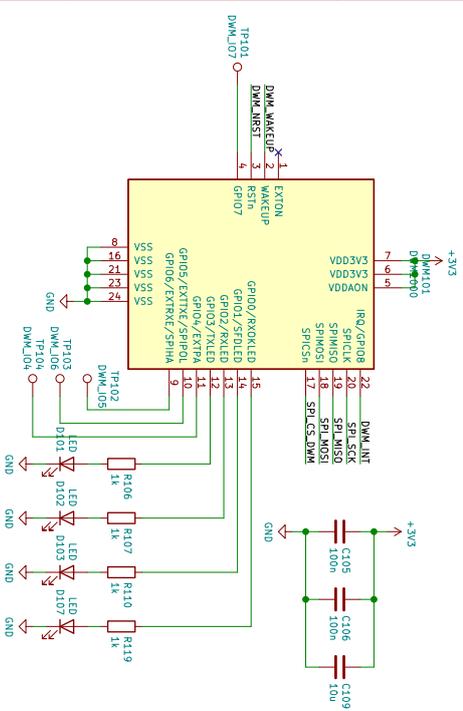
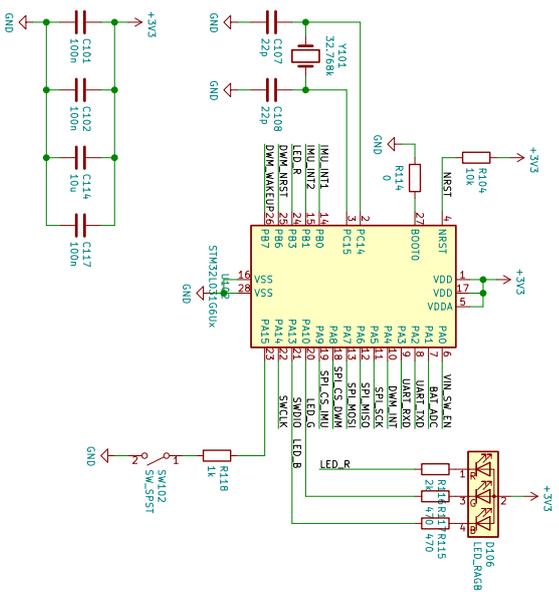
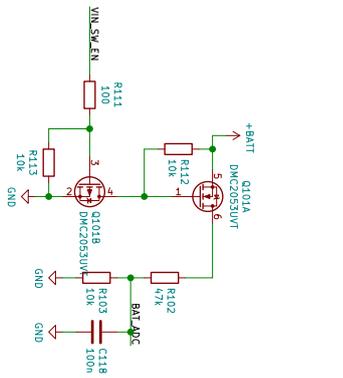
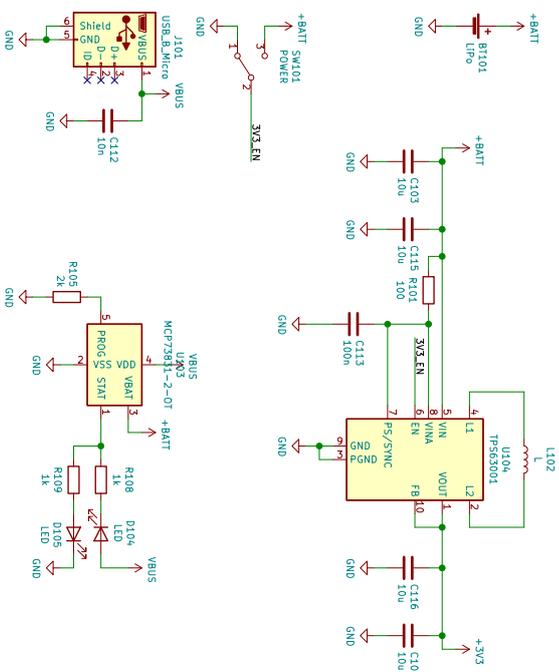
- [1] IEEE Standard for Local and metropolitan area networks–Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs). In: *IEEE Std 802.15.4-2011 (Revision of IEEE Std 802.15.4-2006)* (2011), S. 1–314
- [2] ADVANCED REALTIME TRACKING GMBH & CO. KG: *ARTTRACK5*. – URL <https://ar-tracking.com/en/product-program/arttrack5>. – abgerufen am 05.11.2021
- [3] ALARIFI, Abdulrahman ; AL-SALMAN, AbdulMalik ; ALSALEH, Mansour ; ALNAFESSAH, Ahmad ; AL-HADHRAMI, Suheer ; AL-AMMAR, Mai A. ; AL-KHALIFA, Hend S.: Ultra Wideband Indoor Positioning Technologies: Analysis and Recent Advances. In: *Sensors* 16 (2016), Nr. 5, S. 707
- [4] APPLE INC.: Apple stellt AirTag vor. (2021), April. – URL <https://www.apple.com/de/newsroom/2021/04/apple-introduces-airtag/>. – abgerufen am 21.05.2021
- [5] BALSYS, Rokas: *Asynchronous Actor-Critic Agent*. März 2020. – URL <http://www.pylessons.com/A3C-reinforcement-learning/>. – abgerufen am 24.10.2021
- [6] BOSCH SENSORTEC GMBH: *BMI270 – Data sheet*. 2020. – Revision 1.3
- [7] BUNDESNETZAGENTUR: Frequenzen für Ultra-Wideband-Technologie bereitgestellt. (2008). – URL <https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Allgemeines/Presse/Pressemitteilungen/2008/PM20080116UltraWidebandTechnologieId12430pdf.pdf>. – abgerufen am 17.08.2021
- [8] BUNDESNETZAGENTUR: Allgemeinzuteilung von Frequenzen für die Nutzung durch Ultrabreitbandgeräte (UWB). (2019). – URL <https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/>

- [Telekommunikation/Unternehmen\\_Institutionen/Frequenzen/Allgemeinzuteilungen/FunkanlagenGeringerReichweite/2019Vfg135\\_UWB.pdf](#). – abgerufen am 17.08.2021
- [9] DECAWAVE LTD.: *DECARANGERTLS ARM SOURCE CODE GUIDE*. 2015. – Version 2.2
- [10] DECAWAVE LTD.: *DWM1000 Datasheet*. 2016. – Version 1.8
- [11] DECAWAVE LTD.: *DW1000 USER MANUAL*. 2017. – Version 2.11
- [12] GÄRTNER, Henner ; MAASS, Jochen ; STAHR, Pascal ; GENIS, Miguel M.: *Shared Guide Dog 4.0*. – URL <https://ftz-3i.de/blindenhund/>. – abgerufen am 07.10.2021
- [13] INGRAM, S. J. ; HARMER, D. ; QUINLAN, M.: UltraWideBand Indoor Positioning Systems and their Use in Emergencies. In: *PLANS 2004. Position Location and Navigation Symposium (IEEE Cat. No.04CH37556)*, 2004, S. 706–715
- [14] KASTEN, Markus: Hardwareplattformen für autonome Straßenfahrzeuge im Maßstab 1:87. (2021)
- [15] LABBE, Roger: *Kalman and Bayesian Filters in Python*. <https://github.com/rllabbe/Kalman-and-Bayesian-Filters-in-Python>. 2014
- [16] LANGLEY, Richard B.: Dilution of Precision. In: *GPS World* (1999), Mai, S. 52–59
- [17] MNIH, Volodymyr ; BADIA, Adria P. ; MIRZA, Mehdi ; GRAVES, Alex ; LILICRAP, Timothy ; HARLEY, Tim ; SILVER, David ; KAVUKCUOGLU, Koray: Asynchronous methods for deep reinforcement learning. In: *International conference on machine learning* PMLR (Veranst.), 2016, S. 1928–1937
- [18] PAREIGIS, Stephan ; TIEDEMANN, Tim ; STELLDINGER, Peer ; BECKE, Martin ; MEISEL, Andreas: *autosys*. – URL <https://autosys.informatik.haw-hamburg.de/>. – abgerufen am 03.10.2021
- [19] POUPART, Pascal ; BOUTILIER, Craig: Value-Directed Belief State Approximation for POMDPs. (2013)
- [20] SAHINOGLU, Zafer ; GEZICI, Sinan ; GUVENC, Ismail: *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*. Cambridge University Press, 2008

- [21] SICILIANO, Bruno ; KHATIB, Oussama: *Springer Handbook of Robotics*. Berlin, Heidelberg : Springer-Verlag, 2008
- [22] STMICROELECTRONICS: *STM32L031x4 STM32L031x6*. 2018. – DS10668 Rev 6
- [23] SUTTON, Richard S. ; BARTO, Andrew G.: *Reinforcement Learning: An Introduction*. MIT Press Ltd, 2018
- [24] TANENBAUM, Andrew S. ; WETHERALL, David: *Computer networks, 5th Edition*. Pearson, 2011
- [25] TAYLOR, Robert: *Example Use of the Kalman Filter Algorithm*. April 2018. – URL <http://www.pyrunner.com/weblog/2018/04/12/kalman-example/>. – abgerufen am 05.08.2021
- [26] YAVARI, Mohammadreza ; NICKERSON, Bradford G.: Ultra wideband wireless positioning systems. In: *Dept. Faculty Comput. Sci., Univ. New Brunswick, Fredericton, NB, Canada, Tech. Rep. TR14-230* 40 (2014)
- [27] ZHANG, Ethan ; MASOUD, Neda: Increasing GPS Localization Accuracy With Reinforcement Learning. In: *IEEE Transactions on Intelligent Transportation Systems* 22 (2021), Nr. 5, S. 2615–2626
- [28] ZUCCONI, Alan: *Positioning and Trilateration*. März 2017. – URL <https://www.alanzuconi.com/2017/03/13/positioning-and-trilateration/>. – abgerufen am 04.08.2021

# A Anhang: Schaltplan UWB-Beacon

Quelle: Markus Kasten



# Glossar

**Ground Truth** Als Ground Truth wird eine Information bezeichnet, die als wahr oder real anzusehen ist. Ein Ground Truth wird in der Regel empirisch durch eine direkte Messung oder Beobachtung ermittelt.

**Lidar** Lidar (light detection and ranging) ist ein optisches Verfahren, dass die Entfernung von Objekten bestimmt.

**Odometrie** Die Odometrie bezeichnet die Verwendung von Antriebsdaten zur Schätzung der Positions- und Orientierungsänderung im Laufe der Zeit.

**Paketumlaufzeit** Die Paketumlaufzeit ist die Dauer, die ein Datenpaket benötigt, um vom Startpunkt zum Ziel und wieder zurück zum Startpunkt zu gelangen.

**Radar** Radar (radio detection and ranging) ist ein Verfahren, dass mit elektromagnetischen Wellen die Entfernung von Objekten bestimmt.

**Signal-Rausch-Verhältnis** Das Signal-Rausch-Verhältnis (SNR) ist ein Maß für die Qualität eines Nutzsymbols. Das Signal-Rausch-Verhältnis muss ausreichend groß sein, damit sich das Nutzsymbols vom Hintergrundrauschen abhebt.

**Tiefenkamera** Eine Tiefenkamera bestimmt zusätzlich für jeden Bildpunkt die Distanz des darauf abgebildeten Objektes.

**Transceiver** Ein Transceiver ist eine Kombination eines Senders und eines Empfängers in einem Gerät.

## Erklärung zur selbstständigen Bearbeitung einer Abschlussarbeit

Gemäß der Allgemeinen Prüfungs- und Studienordnung ist zusammen mit der Abschlussarbeit eine schriftliche Erklärung abzugeben, in der der Studierende bestätigt, dass die Abschlussarbeit „— bei einer Gruppenarbeit die entsprechend gekennzeichneten Teile der Arbeit [(§ 18 Abs. 1 APSO-TI-BM bzw. § 21 Abs. 1 APSO-INGI)] — ohne fremde Hilfe selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt wurden. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich zu machen.“

*Quelle: § 16 Abs. 5 APSO-TI-BM bzw. § 15 Abs. 6 APSO-INGI*

## Erklärung zur selbstständigen Bearbeitung der Arbeit

Hiermit versichere ich,

Name: \_\_\_\_\_

Vorname: \_\_\_\_\_

dass ich die vorliegende Bachelorarbeit – bzw. bei einer Gruppenarbeit die entsprechend gekennzeichneten Teile der Arbeit – mit dem Thema:

### **Hochgenaue Positionsbestimmung mit Ultra-wideband und Reinforcement Learning**

ohne fremde Hilfe selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich gemacht.

\_\_\_\_\_  
Ort                      Datum                      Unterschrift im Original