

# Methode für Vorhersagen über die Fortführung von Handbewegungen

Philipp Rall<sup>1</sup>, Nicolas Bender<sup>2</sup>

**Abstract:** Die vorliegende Arbeit befasst sich mit der Entwicklung einer Methode zur Echtzeit-Vorhersage von Trajektorien seitlicher Greifbewegungen zur Kollisionsvermeidung in der kollaborativen Robotik. Ein Neuronales Netz sagt hierfür anhand des Verlaufs der Anfangsbewegung in einem Regressionsansatz die Endposition und Dauer des gesamten Greifvorgangs voraus. Durch das Minimum Jerk Model für gekrümmte Punkt-zu-Punkt-Bewegungen lässt sich daraufhin der weitere Verlauf der Trajektorie präzise berechnen. Die Arbeit legt besonderen Fokus auf die Entwicklung einer automatisierten Pipeline zur Datenvorverarbeitung, die aufgenommene Rohdaten von natürlichen Greifbewegungen in mehreren modularen Verarbeitungsphasen zur qualitativ hochwertigen und vereinheitlichten Trainingsdaten transformiert sowie fehlerbehaftete Messdaten aussortiert.

**Keywords:** Mensch-Maschine Interaktion; Minimum Jerk Model; Datenvorverarbeitung; Kollaborative Robotik; Maschinelles Lernen

## 1 Einleitung

Mit dem Fokus auf innovative Produktionsmethoden und technologische Strategien im Zeichen von Industrie 4.0 werden unter dem Stichwort Human-Robot-Collaboration Szenarien zunehmend realistischer, in denen sich Menschen und Roboter einen gemeinsamen Arbeitsplatz teilen, um kollaborativ Aufgaben bearbeiten zu können.[Ke20] Um diese geteilten Arbeitsplätze für Menschen sicher zu gestalten, ist es unabdingbar, technische Vorsichtsmaßnahmen zu implementieren. Damit soll verhindert werden, dass Menschen von leistungsstarken Robotern verletzt oder empfindliche Roboter von Menschen beschädigt werden.[Ma16]

Eine Möglichkeit eines Kontrollmechanismus ergibt sich aus der Beobachtung der menschlichen Bewegung, die in den Kontext der Bewegungen des Roboters gesetzt werden. Bewegungsvorgänge des Roboters haben den Vorteil gegenüber den menschlichen Bewegungen, dass diese präzise gesteuert, lokalisiert und vorhergesagt werden können. Schwieriger gestaltet sich die Betrachtung der Bewegung des Menschen, die nach von außen nicht zwingend ersichtlichen Einflüssen erfolgt und deshalb kaum von außerhalb kontrolliert werden kann. Die Kontrollmechanismen des Roboters müssen sich daher an die Spontanitäten der menschlichen Bewegungen anpassen und Prognosen für beispielsweise Handbewegungen treffen, um angemessen darauf reagieren zu können.

---

<sup>1</sup> Technische Universität Darmstadt, Karolinenpl. 5, 64289 Darmstadt, mail@philipp-rall.de

<sup>2</sup> Universität Heidelberg, Grabengasse 1, 69117 Heidelberg, nicolas.bender@posteo.de

Das Ziel dieser Arbeit ist die Entwicklung und Bewertung einer Methode, mit der sich die Fortführung von Handgelenken vorhersagen lässt. Dabei soll auf Grundlage einer kurzen Bewegungssequenz der weitere Verlauf dieser Bewegung berechnet werden. Unter Zuhilfenahme von Maschinellem Lernen besteht das Ziel darin, ein Modell zu entwickeln, mit dem aus aufgenommenen Anfangsbewegungsdaten einer Armbewegung die Endposition des Handgelenks und Dauer der Bewegung prognostiziert und die Trajektorie berechnet werden kann.

Abbildung 1 stellt das angestrebte Gesamtszenario der Kollisionsvermeidung im Bereich der kollaborativen Robotik dar. Die Kollisionsberechnung sowie Ermittlung von Handlungs-konsequenzen für den Roboter sind nicht Thema der Arbeit.

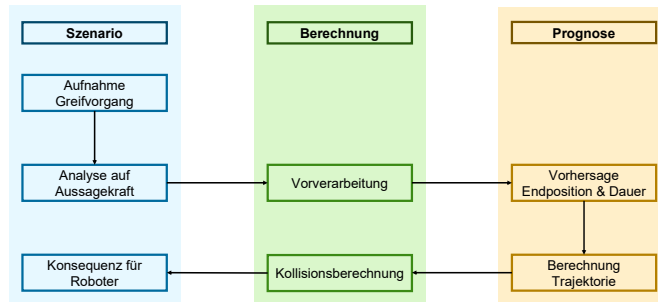


Abb. 1: Angestrebtes Gesamtszenario

## 2 Grundlagen

### 2.1 Experimentelles Setup

Zur Datenaufnahme wird ein Szenario der kollaborativen Robotik mit einem Roboter und einer Testperson simuliert. Ziel des Szenarios ist die Übergabe eines Paketes vom Roboter an den Menschen. Als Roboterarm wird das Modell UR5e von Universal Robots verwendet. Der Roboter positioniert das Paket an einer von 346 vorgegebenen Positionen im Raum. Dazu ist der Raum wie von Rettig et al. beschrieben in Würfel mit einer Kantenlänge von 10cm gleichmäßig aufgeteilt.[Re21] Nach Abschluss der Positionierung kann die Testperson das Paket mit beiden Händen greifen und auf einem Tisch ablegen. Diese zielgerichtete seitliche Greifbewegung ist die für die Datenaufnahme relevante Bewegung.

Die Aufnahme der Bewegungsdaten erfolgt über am Körper der Testperson angebrachte reflektierende Cluster aus Markern und ein 120-Hz-Vicon-Kamerasystem. Dieses besteht aus acht Infrarotkameras, welche das von den Markern reflektierte Infrarotlicht aus mehreren Perspektiven erkennen können. Die Cluster werden entsprechend an den in der Arbeit von Rettig et al. [Re21] sowie Murray et al. [Mu99] genannten Positionen des Torsos angebracht.

## 2.2 Modellierung von Armbewegungen

Die Grundproblematik in der Betrachtung von menschlichen Armbewegungen liegt darin, dass es beliebig viele Möglichkeiten gibt, einen Pfad und die Geschwindigkeit von einem Startpunkt zu einem Endpunkt zu wählen. Für hindernisfreie Bewegungsvorgänge sieht die Literatur die Gemeinsamkeiten für die physikalische Beschreibung in den folgenden vier Aspekten: Ein möglichst glatter und geradliniger Pfad, eine ruckfreie Bewegung, ein unimodales glockenförmiges Geschwindigkeitsprofil und eine Skalierung der Spitzengeschwindigkeit und -beschleunigung sowie deren Zeitpunkte mit der Bewegungsamplitude.[MK99][FH85]

Es wird davon ausgegangen, dass sich die Berechnung eines Bewegungsvorgangs durch das menschliche Gehirn anhand dieser Faktoren annähern lässt. Für den möglichst glatten und geradlinigen Pfad sind Szenarien gemeint, in denen bei einem Greifvorgang ein Endpunkt, eventuell in Form eines Zielobjekts, anvisiert und eine Hand hindernisfrei darauf zu bewegt werden kann. Der Geschwindigkeitsverlauf der Hand ist dabei glockenförmig und enthält in der Spitze eine maximale Geschwindigkeit. Diese Höchstgeschwindigkeit und die maximale Beschleunigung und deren Zeitpunkte sind abhängig von der Entfernung des Endpunkts.

In der Forschung findet sich als etabliertes Modell der Vorhersage von menschlichen Bewegungen das Minimum Jerk Model von Flash und Hogan.[FH85] Diese Methode soll eine mathematische Herangehensweise zur Verfügung stellen, die ohne Einbezug der dahinterliegenden anatomischen Faktoren das Verhalten von willkürlichen menschlichen Bewegungen präzise modellieren kann. Der Grundgedanke dieses Modells ist die Eigenschaft, dass eine menschliche Bewegung einen Bewegungsablauf durchführt, der den kleinsten möglichen Ruck verursacht. Der Ruck ist dabei die dritte Ableitung des Ortes nach der Zeit und physikalisch gesehen damit die Änderung der Beschleunigung. Das Minimum Jerk Model sieht vor, die Trajektorie der Hand bei einem hindernisfreien geradlinigen Greifvorgang zu approximieren und weist als Rahmenbedingung auf, dass die Bewegung aus dem Stillstand startet und sich nur in einer zweidimensionalen Ebene abspielt. Das Modell lässt sich auch für Bewegungen im dreidimensionalen Raum anwenden.[FMv12] Für die Beschreibung eines Bewegungsvorgangs muss das Integral über die Quadrate des Rucks der einzelnen Dimensionen minimiert werden. Somit lässt sich als zu minimierende Kostenfunktion  $C$  in Gleichung 1 definieren:

$$C = \frac{1}{2} \int_0^{t_f} \left( \left( \frac{d^3x}{dt^3} \right)^2 + \left( \frac{d^3y}{dt^3} \right)^2 + \left( \frac{d^3z}{dt^3} \right)^2 \right) dt \quad (1)$$

Durch die Minimierung dieser Kostenfunktion, in der  $x$ ,  $y$  und  $z$  Funktionen über der Zeit  $t$  sind, ergeben sich nach Flash und Hogan [FH85] für verschiedene Arten von Armbewegungen jeweils Funktionen für jede Koordinate, die eine Bewegung mit minimalem Ruck beschreiben. Im Rahmen dieser Arbeit sollen gekrümmte Punkt-zu-Punkt Bewegungen betrachtet werden.

Gekrümmte Punkt-zu-Punkt Bewegungen stellen Bewegungskurven von einem Startpunkt über einen Zwischenpunkt zu einem Zielpunkt dar und können somit einem Hindernis ausweichen. Der Zwischenpunkt sollte dabei die maximale Distanz zur direkten Verbindung von Start- und Endpunkt aufweisen. Zur Beschreibung der Trajektorie ergeben sich durch die Optimierung je Koordinate zwei Funktionen. Eine Funktion für die Punkte vor dem Zwischenpunkt und eine für die Punkte danach.[FH85]

### 2.3 Datenvorverarbeitung

Die Bestimmung der Geschwindigkeit muss bei Messdaten aufgrund von diskreten Werten numerisch erfolgen. Eine Bildung von Vorwärtsdifferenzen ist in der Literatur dabei als ausreichend angesehen.[Di11] Extreme Ausreißer im Geschwindigkeitsprofil können anhand des Ausreißertests nach Grubbs erkannt werden.[Gr50]

Für die Rauschentfernung der aufgenommenen Messdaten nutzen Experimente den Savitzky-Golay Filter.[Zh21] Darüber hinaus werden auch aufwendige Verfahren wie der Kalman Filter verwendet, die zufriedenstellende Ergebnisse für sehr verrauschte Daten liefern können.[AKG04][Ed17]

Um den Beginn einer Bewegung zu definieren, wird von Zhao et al. [Zh21] vorgeschlagen, alles unter 1% der Maximalgeschwindigkeit des Bewegungsvorgangs als Stillstand anzunehmen. Damit lässt sich ausgehend vom Punkt der höchsten Geschwindigkeit zeitlich die Bewegung bis zu den Punkten eingrenzen, an denen der Stillstand erreicht wird.

## 3 Methode

Die vorgeschlagene Methode besteht aus den drei Teilen Datenvorverarbeitung, Vorhersage und der Pfadberechnung.

### 3.1 Datenvorverarbeitung

Die entwickelte Pipeline zur automatisierten Vorverarbeitung von Messdaten natürlicher Handgelenkstrajektorien beinhaltet acht Phasen. Ziel ist die Gewinnung qualitativ hochwertiger und einheitlicher Trainingsdaten.

**Interpolation und Glättung** Zu Beginn der Pipeline werden bis zu drei aufeinanderfolgende Not-a-Number-Werte (NaN) in der Messreihe der Trajektorie durch ein Polynom dritten Grades interpoliert. Treten in den Messdaten längere NaN-Folgen auf, so ist eine qualitativ aussagekräftige Interpolation nicht mehr möglich. In Messdaten vorliegendes Rauschen kann

mit dem Savitzky-Golay Filter geeignet geglättet werden, der auf polynomialer Regression anhand von Stützstellen basiert. Die Verwendung eines Polynoms dritten Grades und 13 Stützstellen liefert für die vorliegenden Datensätze gute Ergebnisse.

**Beschneidung** Die Rohdaten der Armbewegungen bestehen aus unterschiedlich langen Datensätzen, die grob zugeschnitten vorliegen. Die Datensätze enthalten einen Abschnitt zu Beginn, in dem sich die Position der Handgelenke, der Ellenbogengelenke und der Schultergelenke kaum verändern. Dieser Zustand spiegelt das lose Hängen der Arme und Hände der Testperson vor der eigentlichen Greifbewegung wider. Darauf folgend beginnt die Phase des Greifens, in der sich die Position der jeweiligen Gelenke verändert. Da für die Auswertung des Greifvorgangs die Daten ohne Bewegung irrelevant sind, sollen diese aus den Datensätzen ausgeschnitten werden. Der Grundgedanke besteht hierbei darin, dass der Zeitraum des Greifvorgangs über das Geschwindigkeitsprofil der Trajektorie ermittelt werden kann.

Angelehnt an das Vorgehen von Zhao et al. [Zh21] soll zunächst ein Schwellwert auf Basis der Maximalgeschwindigkeit der Trajektorie bestimmt werden, der als Stillstand der Bewegung angenommen wird. Dafür wird der initiale Schwellwert auf 1% der Maximalgeschwindigkeit gesetzt. Ausgehend vom Punkt der höchsten Geschwindigkeit kann anschließend der nächste Punkt vor und nach diesem Zeitpunkt ermittelt werden, zu dem die nach dem Schwellwert als Stillstand angenommene Bewegung auftritt. Damit lässt sich der Datensatz soweit beschneiden, dass der erste bzw. der letzte Punkt des Datensatzes den Beginn bzw. das Ende des Greifvorgangs darstellt.

**Statische Filterung** Diese Phase filtert fehlerhafte Messdaten basierend auf den Kriterien der Länge der Messreihe, Anzahl an NaN-Werten je Messreihe und der Maximalgeschwindigkeit je Messreihe heraus.

Die obere und untere Schranke der Messreihenlänge ergibt sich aus Beobachtungen durchschnittlicher Längen korrekt gemessener Messreihen. Alle Messreihen, die nach der vorangegangenen Interpolation noch NaN-Werte enthalten, besitzen nicht korrigierbare Messfehler und werden aussortiert. Als Maximalgeschwindigkeit für die Bewegung eines menschlichen Handgelenks kann realistisch 5m/s [E114] festgelegt werden.

**Spiegelung und Translation** In einem Teil der vorliegenden Datensätze sind für Greifvorgänge mit beiden Händen die Bewegungspunkte für beide Körperhälften vorhanden, in vielen Fällen jedoch nur für eine. Zur Vereinheitlichung und Vergrößerung der Datenmenge sollen alle Datensätze durch eine Spiegelung auf den Greifvorgang der rechten Körperhälfte abgebildet werden.

Für die Umsetzung der Spiegelung wird folgende Vorgehensweise vorgeschlagen. Liegen für einen Greifvorgang Datensätze zu beiden Körperhälften vor, soll eine Spiegelebene anhand dieser beiden Datensätze ausgerichtet werden. Ist nur eine Körperhälfte vorhanden, wird

die Spiegelung an einer Ebene durchgeführt, die durch Start- und Endpunkt der Bewegung verläuft und senkrecht zur größten Auslenkung der Trajektorie steht.

Nach der Spiegelung erfolgt eine Verschiebung aller Trajektorien, sodass sie jeweils im Ursprung starten.

**Metrikbasierte Filterung** Die weitere Selektion qualitativ hochwertiger Messdaten lässt sich mit Hilfe von drei Metriken und dynamisch gebildeten Grenzwerten durchführen.

1. Handgelenk-Ellenbogen-Distanz: Arithmetisches Mittel der euklidischen Distanzen zwischen Ellenbogen und Handgelenk zu allen Zeitpunkten  $n$ , sollte annähernd konstant sein.
2. Handgelenk-Handgelenk-Distanz: Arithmetisches Mittel der euklidischen Distanzen beider Handgelenke zu allen Zeitpunkten  $n$ , sollte aufgrund der vorhergehenden Spiegelung der Trajektorien der linken Körperhälfte möglichst gering sein.
3. Vergleich der Z-Koordinaten: Arithmetisches Mittel der absoluten Differenzen der z-Koordinaten der Handbewegungen beider Körperseiten zu allen Zeitpunkten  $n$ , sollte nahezu Null sein.

Die Schwellwerte der Metriken, ab welchen ein Messdatensatz als fehlerbehaftet angesehen wird, müssen aufgrund der anatomischen Unterschiede der Testpersonen jeweils anhand der jeweiligen Messdaten ermittelt werden. Der untere und obere Schwellwert  $\kappa_m$  je Metrik  $m$  ergibt sich aus dem Quantil der jeweiligen Metrikwertverteilung, wobei der Unterschreitungsanteil des Quantils abhängig von der Standardabweichung  $s_m$  und einer Skalierungskonstante  $\delta_m$  gewählt wird. Hierdurch ist sichergestellt, dass stark verrauschte Daten stärker aussortiert werden. Für die drei vorgestellten Metriken lassen sich folgende Skalierungskonstanten verwenden:  $\delta_1 = 0.009$ ,  $\delta_2 = 0.005$ ,  $\delta_3 = 0.008$

$$\kappa_m^{unten} = Q(\delta_m s_m) \quad \text{mit } \delta_m s_m \in [0, 0.5] \quad (2)$$

$$\kappa_m^{oben} = Q(1 - \delta_m s_m) \quad \text{mit } \delta_m s_m \in [0, 0.5] \quad (3)$$

Nach der Berechnung der Schwellwerte je Metrik und Messreihe erfolgt in einem mehrstufigen Verfahren die Entscheidung über die Aussortierung eines Datensatzes. Hierzu wird bei auffälligen Metrikwerten außerhalb der Schwellwerte der normierte Abstand zum Mittelwert der jeweiligen Verteilung als Ausreißerintensität  $\rho_m$  herangezogen. Das Aussortierungsverfahren bildet anschließend wie in Algorithmus 1 eine gewichtete Summe der Ausreißerintensitäten je Metrik und vergleicht diese mit einer Fehlerschranke.

Die Wahl der Gewichte spiegelt die Relevanz der einzelnen Metriken wider. Höchste Relevanz haben die Ergebnisse der Handgelenk-Handgelenk-Metrik, da Ausreißer in dieser Metrik klar auf Fehler in den für das maschinelle Lernen relevanten Handgelenkdaten zurückzuführen sind.

```

für alle translatierten Dateien tue
   $\epsilon \leftarrow 0$ ;
  wenn Datei in Handgelenk-Ellenbogen-Metrik als Ausreißer markiert dann
     $\epsilon \leftarrow \epsilon + 2 \cdot (1 + \text{Ausreißerstärke } \rho_1 \text{ der Datei})$ ;
  Ende
  wenn Datei in Handgelenk-Handgelenk-Metrik als Ausreißer markiert dann
     $\epsilon \leftarrow \epsilon + 4 \cdot (1 + \text{Ausreißerstärke } \rho_2 \text{ der Datei})$ ;
  Ende
  wenn Datei in Z-Differenz-Metrik als Ausreißer markiert dann
     $\epsilon \leftarrow \epsilon + 2 \cdot (1 + \text{Ausreißerstärke } \rho_3 \text{ der Datei})$ ;
  Ende
  wenn  $\epsilon > 4$  dann
    Sortiere Datei aus;
  Ende
Ende

```

**Algorithmus 1:** Gewichtete Aussortierung fehlerhafter Datensätze

**Rotation** Diese Phase rotiert die Handgelenkstrajektorien, sodass die Anfangsverläufe der Bewegungen einheitlich in  $x$ -Achsenrichtung zeigen. Der Anfangsverlauf ist hierbei die feste Anzahl an Datenpunkten, die als Eingang für das zu trainierende Neuronale Netz gilt. Der letzte Datenpunkt des Anfangsverlaufes lässt sich als Rotationspunkt bezeichnen, dessen Ortsvektor die Richtung des Anfangsverlaufs darstellt. Für die Rotation gilt der Drehwinkel  $\psi$  zwischen der  $x$ -Achse und dem Ortsvektor  $\vec{d}$  des Rotationspunktes. Die Rotationsachse stellt die Normale der von der  $x$ -Achse und  $\vec{d}$  aufgespannten Ebene dar. Die rotierten Daten ergeben sich durch Multiplikation mit einer entsprechenden Rotationsmatrix.

### 3.2 Vorhersage

Die Vorhersage der zur Pfadberechnung relevanter Angaben der Endposition und Dauer der Handgelenksbewegung erfolgt über zwei separate Neuronale Netze. Als Eingangsmerkmale lassen sich die drei Koordinaten der Zeitpunkte der zuvor in der Rotation definierten Anfangsbewegung verwenden. Diese hat im Beispiel eine Länge von 85 Zeitpunkten, d.h. es ergeben sich 255 Eingangsmerkmale.

Für die Vorhersage der Endposition hat sich empirisch ein Neuronales Netz mit Dense-Layern als sinnvoll erwiesen, siehe dazu Abbildung 2. In den verborgenen Schichten verwendet das Netz als Aktivierungsfunktion Rectified Linear Units (ReLUs) und in der Ausgabeschicht eine lineare Aktivierungsfunktion. Die Fragezeichen in der Abbildung stehen für eine beliebige Anzahl an Datensätzen. Vor dem Training erfolgt eine Skalierung der Daten in den einheitlichen Wertebereich von 0 bis 1. Als Kostenfunktion wird zur Vorhersage der Endpunkte die quadrierte euklidische Distanz zwischen dem vorhergesagten und dem tatsächlichen Endpunkt verwendet. Für die Vorhersage der Dauern erweist sich der Mean Absolute Error als geeignete Kostenfunktion, da keine räumlich interpretierbaren Daten vorliegen.

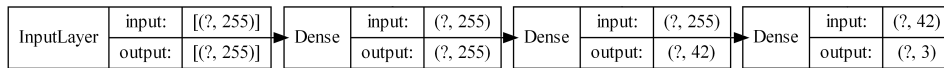


Abb. 2: Architektur des neuronalen Netzes zur Vorhersage der Endpositionen

Das Training des Neuronalen Netzes geschieht über Datensätze, die nach der in der vorliegenden Arbeit ausgeführten Methode vorverarbeitet sind. Die Datensätze bestehen aus Greifbewegungen von verschiedenen Probanden, die nach dem in Abschnitt 2.1 genannten Szenario aufgenommen sind.

### 3.3 Pfadberechnung

Mittels der in diesem Kapitel prognostizierbaren Endpunkte und Bewegungsdauern wird mit dem Minimum Jerk Modell für gekrümmte Bewegungen der Verlauf der Trajektorie anhand der aufgenommenen Anfangsbewegung vorausberechnet. Notwendig hierfür ist neben der Information des Start- und des Endpunkts  $P_1$  und  $P_2$  der Bewegung sowie der Dauer des Greifvorgangs auch die Angabe eines dritten Punkts  $P_C$ , durch den die Bahnkurve verlaufen soll. In der idealen Modellierung der realen Trajektorie mit diesem Verfahren ist dies der Punkt der größten Entfernung von der Strecke zwischen Start- und Endpunkt. Da dieser bei der Betrachtung der Anfangsbewegung noch nicht aufgetreten sein muss, soll die Annahme getroffen werden, dass derjenige Punkt der Anfangsbewegung mit dem größten Abstand von der Strecke zwischen Startpunkt und dem durch das ML-Modell prognostizierten Endpunkt diesen dritten Punkt  $P_C$  darstellt. Das Vorgehen zur Bestimmung der gekrümmten Trajektorie sieht vor, eine Berechnung der Bahnkurve zeitlich vor und nach dem Punkt  $P_C$  durchzuführen.

## 4 Ergebnisse

Als bestmögliche Konfiguration erweist sich eine Anfangsbewegung der Länge 85, somit werden durchschnittlich 65% der Trajektorie als Eingangsdaten zur Vorhersage der Endposition benötigt. Der Rotationspunkt, nach dem sich die Rotation ausrichtet, liegt dabei beim 70. Punkt der Trajektorie. Als Epochenanzahl des Trainings stellt sich 350 als sinnvolle Größe heraus, da bei dieser Epochenanzahl das Neuronale Netz bei Testdaten ähnlich gute Ergebnisse wie bei Trainingsdaten erzielt und noch keine starke Überanpassung auf die Trainingsdaten vorliegt. Als Batch-Size lässt sich 250 festlegen, von den Trainingsdaten werden 25% als Testdaten und 5% als Validierungsdaten verwendet. Bezüglich der Metrikergebnisse können bei bereits beschriebener Trainingskonfiguration die Endpunkte durchschnittlich bei den Trainingsdaten auf 10.7cm und bei den Testdaten auf 13.7cm genau vorhergesagt werden, die Werte ergeben sich als durchschnittliche euklidische Distanz der letztendlich vorhergesagten und tatsächlichen Endpunkte. Bewegungsdauern können auf bis zu zehn Zeitpunkte genau bei Trainings- und Testdaten vorhergesagt werden.



Die Berechnung der Bahnkurve durch das Minimum Jerk Model für gekrümmte Punkt-zu-Punkt Bewegungen anhand der Anfangsbewegung und dem vorhergesagten Endpunkt und Dauer ergibt einen ähnlichen Verlauf im Vergleich zur tatsächlichen Trajektorie. In Abbildung 3 ist eine beispielhafte Gegenüberstellung zu sehen. Abgebildet ist zum einen der Verlauf der im Versuch aufgenommenen Trajektorie in blau und zum anderen die anhand dieser Daten und der Vorhersagen des Neuronalen Netzes berechneten Bahnkurve in orange.

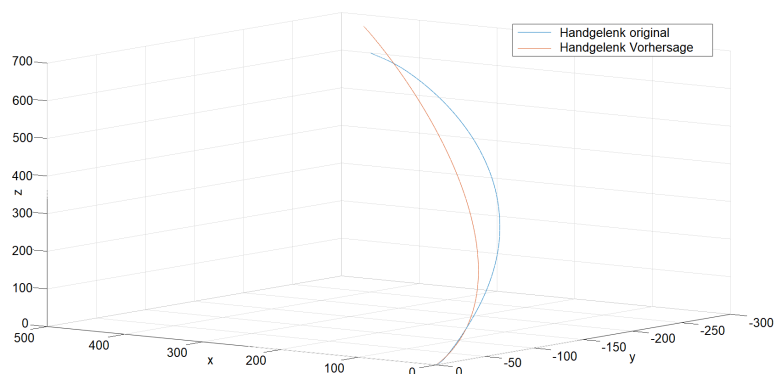


Abb. 3: Vergleich der tatsächlichen (blau) zur berechneten Trajektorie (orange)

## 5 Diskussion

**Datenvorverarbeitung** In Bezug auf die Datenvorverarbeitung lässt sich in der Phase der Beschneidung kritisieren, dass sich die Bestimmung der Qualität der Beschneidung lediglich auf empirisch bestimmte Werte begründet. Eine Aussage über die tatsächliche Güte der Datensätze nach der Beschneidung findet nur durch eine manuelle Analyse der Geschwindigkeitsverläufe und der Trajektorien statt. Diese entspricht allerdings nicht dem Gedanken der vollständigen Automatisierung der Datenvorverarbeitung und ist mit erheblichem Aufwand verbunden.

Durch stichprobenartig visualisierte Datensätze fällt auf, dass sich fehlerbehaftete Daten vor allem durch Sprünge in den Geschwindigkeitsprofilen, harte Kanten oder sonstige ungewöhnliche Form auszeichnen. Die Filterung deckt diese Kriterien bisher nicht vollständig ab, beispielsweise wäre die Ergänzung um eine Stetigkeitsanalyse der Geschwindigkeitsverläufe sinnvoll.

**Machine Learning** Es lässt sich der generelle Erfolg des Maschinellen Lernens diskutieren, da fast zwei Drittel der Trajektorie als Eingangsdaten benötigt werden, um den Endpunkt und die Dauer vorherzusagen. Zudem ist eine Vorhersage der Dauer mit einer Präzision von 10 Zeitpunkten unter der Betrachtung, dass aufgrund der statischen Filterung noch

maximal 90 Zeitpunkte, durchschnittlich jedoch lediglich 45 weitere Messpunkte folgen, relativ unpräzise. Da der Fokus dieser Arbeit allerdings auf der Datenvorverarbeitung liegt, ist dieses Ergebnis akzeptabel. Es lassen sich trotz verbesserungsfähiger Präzision gute Ergebnisse in der Berechnung der weiteren Trajektorie erzielen. Zudem gilt zu beachten, dass in der Praxis zur Kollisionsvermeidung ohnehin ein Sicherheitsbereich um die berechnete Trajektorie gebildet werden muss. Dadurch lassen sich Ungenauigkeiten in der Berechnung etwas ausgleichen.

**Praxistauglichkeit** Für die Anwendung in der Praxis ist eine Vorverarbeitung der Daten der Anfangstrajektorie erforderlich. Durchschnittlich folgen auf den 85. Messwert noch ca. 45 weitere, bis die Bewegung die Endposition erreicht. Bei einem verwendeten Kamerasystem mit 120Hz bleiben zur Datenverarbeitung und Endpunktvorhersage lediglich 375ms Zeit. Deshalb müssen die Phasen der Datenvorverarbeitung priorisiert werden. Höchste Priorität haben dabei Phasen, die der Vereinheitlichung der Daten dienen, wie die Translation auf den Ursprung und Rotation der Anfangsrichtung in x-Achsenrichtung, da ohne diese Phasen die gemessenen Eingangsdaten nicht mit den Trainingsdaten vergleichbar sind. In Priorität zwei sind Verarbeitungsschritte, die lediglich die Datenqualität steigern. Darunter fallen beispielsweise Interpolation, Glättung und Filterung der Daten.

In Tests können die beschriebenen Verarbeitungsschritte der höchsten Priorität, die Vorhersage des Endpunkts und der Dauer sowie anschließende Retransformation in durchschnittlich 52ms durchgeführt werden. Die Berechnung der vorherzusagenden Bahnkurve der Hand über das Minimum Jerk Model nimmt in Tests etwa 250ms in Anspruch. Über eine numerische Optimierung der Gleichung und der Berechnung sowie Approximationen, wie beispielsweise eine Betrachtung jedes zehnten Frames, ließe sich die Berechnungszeit auf schätzungsweise ein Zehntel davon minimieren. Die gesamte Vorausberechnung der Trajektorie der Hand liegt damit ungefähr bei 80ms und ist somit praxistauglich.

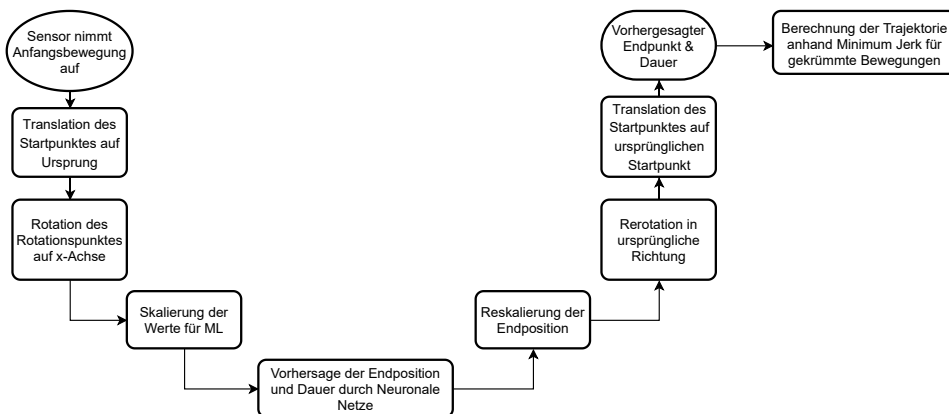


Abb. 4: Ablauf zur Anwendung in der Praxis

Es lässt sich der generelle Ansatz der Kollisionsvermeidung durch Vorhersage der Bewegungsabläufe des Menschen und anschließenden Korrektur der Roboterbewegung kritisch betrachten. Der Ansatz basiert auf der Annahme, dass der Mensch seine Handbewegung unabhängig von Einflüssen der Umwelt deterministisch durchführt. Menschen reagieren allerdings auf äußere Veränderungen, wie beispielsweise eine Routenanpassung oder Pausierung der Roboterbewegung, und können daraufhin ihre Bewegung unmittelbar auch ändern, sodass sie nicht mehr der vorausgerechneten Trajektorie entspricht. Dieser Rückkopplungseffekt muss bei weiteren Überlegungen und dem praktischen Einsatz des Verfahrens bedacht werden.

## 6 Fazit und Ausblick

Zusammenfassend zeigt die vorliegende Arbeit ein Gesamtkonzept zur Vorausberechnung von Handbewegungen auf. Hierbei deckt die Arbeit von der Datenaufnahme, über die Datenvorverarbeitung und dem letztendlichen Training eines Neuronalen Netzes zur Vorhersage von Endpunkt und Dauer von Handbewegungen alle wichtigen Phasen des Maschinellen Lernens ab. Dieser Prozess wird zudem durch die anschließende mathematische Berechnung der Trajektorie in das praktische Anwendungsszenario der Kollisionsvermeidung in der kollaborativen Robotik eingebettet.

Mögliche Erweiterungen und Fortführungen bestehen zum einen in einer weiteren Verbesserung der Datenvorverarbeitung durch beispielsweise eine Erweiterung der Filterung um die Glattheit der Kurven. Grundsätzlich bietet sich auch Maschinelles Lernen als Unterstützung zur Aussortierung von fehlerbehafteten Rohdaten an. Anhand der durch die Pipeline erzeugten gelabelten Daten ließe sich in einem überwachten Lernverfahren zu diesem Zweck ein Klassifikator trainieren.

Zum anderen kann über eine Diskretisierung des Raums in Würfel mit Kantenlänge von 10cm und diskrete Vorhersage der Endposition nachgedacht werden. Aufgrund des erforderlichen Sicherheitsabstandes ist für die Praxis eine exakte kontinuierliche vorhergesagte Position nicht erforderlich.

Auch die Verwendung von Recurrent Neural Networks oder Convolutional Neural Networks lässt sich als Alternative zu den verwendeten Dense-Layern zur Vorhersage der Endpositionen heranziehen.

Für die Anwendbarkeit in der Praxis ist eine Umsetzung ohne markerbasierte Handgelenkerkennung durch ein Kamerasystem und Software wie *OpenPose*, einem Open-Source-Echtzeitsystem zur 2D-Positionserkennung von anatomischen Schlüsselpunkten wie beispielsweise Handgelenken denkbar.[Ca21] Die 2D-Positionsdaten müssten beim aktuell trainierten Modell allerdings durch mehrere Kameras in 3D-Positionsdaten umgerechnet werden.

Darüber hinaus kann die Möglichkeit der Aufnahme von Handbewegungen über eine Kinect-Kamera wie von Elgendi et al. [El14] beschrieben für eine bessere Praxistauglichkeit weiter verfolgt werden.

## Literaturverzeichnis

- [AKG04] Admiraal, Marjan A.; Kusters, Martijn J. M. A. M.; Gielen, Stan C. A. M.: Modeling kinematics and dynamics of human arm movements. 8(3):312–338, 2004. Journal Article.
- [Ca21] Cao, Zhe; Hidalgo, Gines; Simon, Tomas; Wei, Shih-En; Sheikh, Yaser: OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. 43(1):172–186, 2021. Journal Article.
- [Di11] Ding, Hao; Reissig, Gunther; Wijaya, Kurniawan; Bortot, Dino; Bengler, Klaus; Stursberg, Olaf: Human arm motion modeling and long-term prediction for safe and efficient Human-Robot-Interaction. In: 2011 IEEE International Conference on Robotics and Automation (ICRA). IEEE / Institute of Electrical and Electronics Engineers Incorporated, S. 5875–5880, 2011.
- [Ed17] Edvarsson, Andreas: Online Predictions of Human Motion. 2017.
- [El14] Elgendi, Mohamed; Picon, Flavien; Magnenat-Thalmann, Nadia; Abbott, Derek: Arm movement speed assessment via a Kinect camera: a preliminary study in healthy subjects. 13:88, 2014. Journal Article Research Support, Non-U.S. Gov't.
- [FH85] Flash, Tamar; Hogan, Neville: The coordination of arm movements: an experimentally confirmed mathematical model. 5(7):1688–1703, 1985.
- [FMv12] Fligge, Nadine; McIntyre, Joseph; van der Smagt, Patrick: Minimum jerk for human catching movements in 3D. In (Author, Ieee Corporate, Hrsg.): 2012 4th IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics. IEEE, S. 581–586, 6/24/2012 - 6/27/2012.
- [Gr50] Grubbs, Frank E.: Sample Criteria for Testing Outlying Observations. 21(1):27–58, 1950.
- [Ke20] Keibel, Andreas: Mensch-Roboter-Kollaboration in der Medizin. In (Buxbaum, Hans-Jürgen, Hrsg.): Mensch-Roboter-Kollaboration, S. 133–143. Springer Fachmedien Wiesbaden, 2020.
- [Ma16] Markis, Alexandra; Montenegro, Harald; Neuhold, Michael; Oberweger, Andreas; Schlosser, Christian; Schwald, Christoph; Sihm, Wilfried; Ranz, Fabian; Edtmayr, Thomas; Hold, Philipp et al.: Sicherheit in der Mensch-Roboter-Kollaboration. 2016.
- [MK99] Messier, Julie; Kalaska, John F.: Comparison of variability of initial kinematics and endpoints of reaching movements. 125(2):139–152, 1999. Comparative Study Journal Article Comparative Study Journal Article.
- [Mu99] Murray, Ingram Andrew: Determining upper limb kinematics and dynamics during everyday tasks. 1999.
- [Re21] Rettig, Oliver; Müller, Silvan; Strand, Marcus: Determination of posture comfort zones for robot-human handover tasks. 2021.
- [Zh21] Zhao, Jing; Gong, Shiqiu; Xie, Biyun; Duan, Yaxing; Zhang, Ziqiang: Human arm motion prediction in human-robot interaction based on a modified minimum jerk model. 35(3-4):205–218, 2021.