
Machine-Learning-Ansatz zur Bestimmung der Target-Position eines optischen Messgeräts

Luis GARCIA, Uwe BIELKE, Cornelius NEUMANN, Theo HAGENEY,
Klaus BANZHAF, Ernst WIEDENMANN und Rainer BÖRRET

Zusammenfassung

In diesem Beitrag wird ein Machine-Learning-Ansatz zur Bestimmung der Position eines Targets vorgestellt, wie es in einem neuartigen Messgerät namens Multi-Aperture-Positioning-System (MAPS) verwendet wird. Das MAPS besteht aus zwei Komponenten, der Messeinheit selbst und dem LED-Target, welches an einer zu prüfenden Maschine, zum Beispiel einem Industrieroboter oder einem Koordinatenmessgerät (KMG), angebracht ist. Das Messsystem basiert auf dem Prinzip der Lochkamera. Durch Verwendung einer Aperturmaske kommen bei diesem Messgerät mehrere hundert Lochkameras zum Einsatz. Dies ist beschrieben durch (BIELKE et al. 2022b). Um eine hohe Genauigkeit bei der Positionsberechnung zu erreichen, werden mehrere komplexe Algorithmen mit hohem Rechenaufwand eingesetzt. Aktuell bietet das MAPS eine Genauigkeit, die gleich oder besser ist als bei bestehenden photogrammetrischen Messgeräten.

Ein Neuronales Netz (NN) wird anhand von simulierten Daten trainiert und validiert. Die Performance wird mit denen der bisher verwendeten Algorithmen verglichen. Um die Datensätze zu generieren, wird eine proprietäre Simulationssoftware verwendet, die die künstlichen Bilder erzeugt. Bei den Bildern handelt es sich um die Abbildung eines Lichtspots welche entsteht, wenn das Licht des LED-Targets durch die Aperturmaske auf den Kamerasensor scheint. Simuliert werden Lichtspots mit variierendem Schwerpunkt im Subpixelbereich des Sensors.

Der Grundgedanke hinter dem Machine-Learning-Ansatz ist die Untersuchung, ob ein neuronales Netz genauso gut oder besser performen kann wie die bisher verwendeten Algorithmen. Darüber hinaus soll untersucht werden, ob systematische Fehler und Modellfehler reduziert werden können. Während das Neuronale Netz direkt aus den erfassten Daten lernt, verwenden Algorithmen Formeln um den Schwerpunkt der Spots zu berechnen. Diese Formeln haben viele Einflussfaktoren, die nicht immer genau bestimmt werden können, sich mit der Zeit ändern oder gar nicht bekannt sind. Bei der Verwendung eines Neuronalen Netz hingegen lernt das Modell die Beziehung zwischen Eingabe- und Ausgabedaten unter Einbeziehung aller zum Trainingszeitpunkt bestehenden Einflüsse. Insgesamt wird in diesem Beitrag untersucht, ob ein Neuronales Netz die ursprünglich verwendeten Algorithmen ersetzen kann und dabei eine ähnliche oder bessere Leistung erzielt.

1 Einführung

Werkzeugmaschinen, Roboter und Messmaschinen sind heute in fast allen großen Industriezweigen wie der Automobilindustrie, der Elektronikfertigung oder der Produktion von Komponenten und Konsumgütern zu finden. Da die Nachfrage nach individualisierten Produkten steigt, und die Produktionsserien schrumpfen, gewinnen Mehrzweckmaschinen immer mehr an Bedeutung (SCHWENKE et al. 2008). Die Fähigkeit, genaue Teile herzustellen, ist eines der wichtigsten Leistungskriterien für die moderne Produktion. Daher gewinnen Messmaschinen als Hauptinstrument zur Prüfung der Maßhaltigkeit gefertigter Teile zunehmend an Bedeutung. Insbesondere dort, wo eine hohe Wiederholgenauigkeit erforderlich ist, sind Messmaschinen mit hoher Leistungsfähigkeit unverzichtbar. Durch die Forderung nach höherer Produktqualität ist die Verbesserung der Messgenauigkeit zu einem äußerst wichtigen Forschungsbereich geworden (HUANG & NI 1995). Eine Erhöhung der Maschinengenauigkeit kann wiederum durch eine Korrektur erfolgen. Entweder durch Korrektur in einer Rückkopplungsschleife (KUNZMANN et al. 2005) oder durch Kalibrierung unter Anwendung einer Korrekturmatrix (SARTORI & ZHANG 1995).

Ein weiterer wichtiger Punkt in der modernen Produktion ist die Koordination mehrerer Maschinen, die gemeinsam an einer Aufgabe oder einem Werkstück arbeiten. Hier ist die Synchronisierung der Komponenten der entscheidende Teil. Um dies sicherzustellen, müssen alle am Prozess beteiligten Maschinen und Werkstücke präzise und in Echtzeit vermessen werden. Aktuell verfügbare Messgeräte, wie der Laser Tracker oder kamerabasierte Systeme, erfüllen nicht alle der folgenden Kriterien zugleich: Online-Messung, hohe Genauigkeit und gleichzeitige Messung mehrerer Koordinatensysteme. Deshalb wurde MAPS, ein neues photogrammetrisches Messsystem, entwickelt. Es besteht im Wesentlichen aus einem Kamerasensor, einer Aperturmaske und einem LED-Target.

Dieser Beitrag konzentriert sich auf die Entwicklung einer neuen Methode zur Bestimmung des Schwerpunkts der Lichtpunkte, die von MAPS zur Berechnung der Position des LED-Targets verwendet werden. Das Ziel ist es, die traditionell verwendeten Algorithmen durch ein bildregressives Neuronales Netz zu ersetzen. Diese haben sich bereits in einem breiten Spektrum von Anwendungen bewährt. Von der Erkennung und Zählung von Fahrzeugen (TAYARA et al. 2017) bis zur Quantifizierung von Cyanobakterien (PYO et al. 2019) oder der Vorhersage des Alters von Kinderknochen (REN et al. 2019). Neuronale Netze sind äußerst vielseitig, da sie durch Funktionsannäherung selbstständig lernen Beispiele von Eingangsdaten bestmöglich auf Beispiele von Ausgangsdaten abzubilden (IAN et al. 2016).

Zuerst wird die Methodik vorgestellt und die Grundlagen, die für diese Arbeit notwendig sind. Dazu gehören die notwendigen Schritte zur Berechnung der Position des Targets mit MAPS. Die Simulation der Lichtpunkte und das Training des Neuronalen Netzes durch diese ist der Hauptteil. Es wird auch erklärt, wie das NN aufgebaut und implementiert wird. Die Anwendung des NN zur Vorhersage des Zentrums der Lichtpunkte in einem Experiment demonstriert seine Performance, die abschließend mit den traditionell verwendeten Algorithmen verglichen wird.

2 Methodik

Dieser Abschnitt Beschreibt vorangehend welche Algorithmen in MAPS genutzt werden um die Target-Position zu berechnen. Es werden die Lichtspots beschrieben, die in den Messaufnahmen von MAPS zu finden sind und durch welche Methoden man deren Zentrum berechnen kann. Anschließend wird eine Simulation dieser Lichtspots vorgestellt und ein neuer Ansatz das Zentrum dieser mit einem Neuronalen Netz zu bestimmen.

2.1 Bestimmung der Target-Position durch MAPS

Um die Position der Lichtquelle zu bestimmen, verwendet MAPS eine Abfolge verschiedenen Algorithmen, mit denen das vom Kamerasensor aufgenommene Bild verarbeitet wird. Ein Bild stellt eine Matrix von Lichtpunkten dar. Diese werden durch das Licht des LED-Targets abgebildet, das durch die Löcher der Aperturmaske auf den Kamerasensor fällt. Um die Position des LED-Targets aus diesen Bildern zu berechnen, sind die folgenden Schritte erforderlich (BIELKE et al. 2022b):

1. Im ersten Schritt wird ein Unschärfe- und Schwellenwertfilter auf das Bild angewandt, der es ermöglicht, die Lichtpunkte vom Rest des Bildes zu unterscheiden.
2. Danach wird ein Peak-Finding-Algorithmus verwendet, um das Zentrum jedes einzelnen Lichtspots zu finden. Derzeit kann dafür ein Gauß-Fit-Algorithmus (HAGEN & DERENIAK 2008) oder die Methode der Momente (BOWMAN & SHENTON 1998) verwendet werden.
3. Anschließend identifiziert das Programm die Marker im Bild, anhand derer die Position auf der Aperturmaske bestimmt wird. Dies ist notwendig, um im nächsten Schritt die Lichtpunkte mit den Blendenlöchern zu kombinieren.
4. Schließlich wird ein Vektor vom Zentrum jedes Lichtpunkts durch das entsprechende Loch in der Aperturmaske definiert (insgesamt etwa 700), woraufhin deren Schnittpunkt berechnet und gemittelt werden. Der resultierende Schnittpunkt entspricht der Position der LED im MAPS Koordinatensystem.

Die größte Herausforderung besteht darin, den Schwerpunkt (Zentrum) der Lichtpunkte zu bestimmen. Da es nicht ausreicht, diesen auf einen Pixel genau zu berechnen, müssen Methoden verwendet werden, die das Zentrum im Subpixelbereich bestimmen können. Beim derzeitigen Stand der Technik werden hierfür zwei verschiedene Algorithmen verwendet. Während der Gauß-Algorithmus genauer ist, ist der Moments-Algorithmus schneller. Die Genauigkeit bei der Bestimmung der Position des Spotzentrums wirkt sich direkt auf die Genauigkeit bei der Bestimmung der LED-Position aus. Aus diesem Grund wird ein Algorithmus gefordert, der so genau wie möglich ist. Eine möglichst kurze Rechenzeit ist ebenfalls wünschenswert um eine möglichst hohe Messfrequenz zu erreichen. Aus diesem Grund wird in dieser Arbeit ein neuer Ansatz zur Bestimmung der Schwerpunktposition, mit einem Neuronalen Netz, vorgestellt.

2.2 Eigenschaften der realen Lichtspots

Im Folgenden wird ein vom Kamerasystem aufgenommenes Bild in seine einzelnen Lichtspots aufgeteilt und untersucht. In Abbildung 1 (links) ist einer dieser Lichtspots, sowie dessen horizontaler Querschnitt (rechts), zu sehen. Die Größe eines Lichtspots beträgt ca. $100 \times$

100 Pixel, welche durch den Durchmesser der Apertur, die Kammerkonstante und die Größe bzw. Auflösung des Kamerasensors gegeben ist.

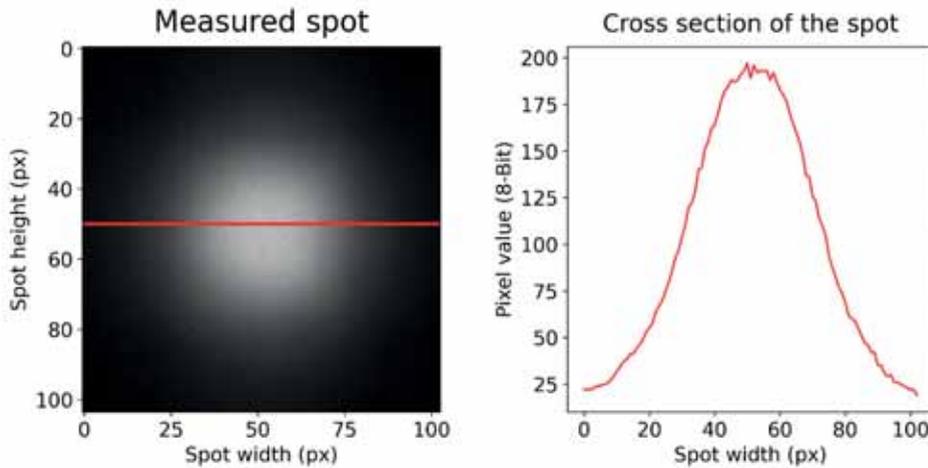


Abb. 1: Reale Abbildung eines Lichtspots auf dem Kamerasensor gemessen durch MAPS (links). Horizontaler Querschnitt des Spots durch den Mittelpunkt (rechts). Auf der y-Achse ist der 8-Bit-Wert der einzelnen Pixel aufgetragen, was einem Wertebereich von [0, 255] entspricht.

Im Querschnitt lässt sich eine gaußähnliche Normalverteilung des Lichts erkennen, die deutlicher wird durch eine Mittelung aller Spots in einer Messaufnahme. Darüber hinaus kann das natürliche Sensorrauschen beobachtet werden, welches sich durch Verzerrungen auf der Linie abbildet. Es beträgt ca. 4 % des maximalen Pixelwerts (255) pro Pixel, wie in BIELKE et al. (2022a) beschrieben. Die hier ermittelten Eigenschaften der Spots dienen als Parameter für die nachfolgende Simulation.

2.3 Simulation der Lichtspots

Die Simulation der Lichtspots ist ein entscheidender Teil dieser Arbeit. Sie ermöglicht die Generierung von umfangreichen Datensätzen, die für das Training des Neuronalen Netz genutzt werden. Die künstlich erzeugten Lichtspots sollten den realen so nahe wie möglich kommen, um später eine gute Leistung des NN zu gewährleisten. Die Simulation wird in einem eigens entwickelten Python-Programm realisiert, wobei unterschiedliche Parameter der Simulation eingestellt werden können. Darunter zählen das natürliche Sensorrauschen, die Spotgröße, sowie verschiedene Filter (Gauß und Fourier) und am wichtigsten – die Position des Spotmittelpunkts. Die Simulation wurde von BIELKE et al. (2022) in „*Simulation and accuracy evaluation of a new 3D photogrammetric position measurement system*“ entwickelt und validiert.

Die Simulation basiert auf der zweidimensionalen Gauß’schen Normalverteilung, welche durch Formel (1) definiert ist (SQUIRES 2001):

$$f(x, y) = A \exp\left(-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2}\right)\right), \quad (1)$$

wobei x_0 und y_0 die Koordinaten des Spotmittelpunkts repräsentieren.

Prinzipiell werden innerhalb eines Datensatzes die Parameter nicht geändert. Das simulierte Sensorrauschen ist zufällig (0 – 6 %), wodurch jedes Bild verschieden verrauscht ist. Das Rauschen soll verschiedene Ausprägungen des natürlichen Sensorrauschens nachahmen. Innerhalb eines Datensatzes variiert nur die Position des Spotmittelpunkts in jedem Spot. Diese wird um Sub-Pixel um den Ursprung (Mittelpunkt des Bildes) verschoben. Um die Datenmenge zu begrenzen, wird der Spotmittelpunkt nur in positive x-Richtung verschoben und nur um maximal ein Pixel. Da die simulierten Spots achsensymmetrisch sind, soll dies vorläufig als *proof-of-concept* dienen. Datensätze werden entweder erzeugt durch festlegen einer Schrittweite, die dann über die Anzahl der Bilder entscheidet. Alternativ kann das Spotzentrum für jedes Bild zufällig gewählt werden.

In Abbildung 2 ist ein Spot mit Zentrum im Ursprung (links) und ein verschobenes Zentrum (rechts) dargestellt. Die Heatmap zeigt einen zentralen Ausschnitt des Spots bei einer Größe von 10×10 Pixeln. Die numerischen Werte innerhalb der Kästchen stellen die Intensität der einzelnen Pixel dar. In der Heatmap links haben die vier zentralen Pixel die maximale Sättigung, somit liegt das Spotzentrum in deren Mitte. Rechts sind nur noch zwei Pixel in maximaler Sättigung, dass daraus resultierende Spotzentrum liegt 0,5 Pixel rechts vom Ursprung.

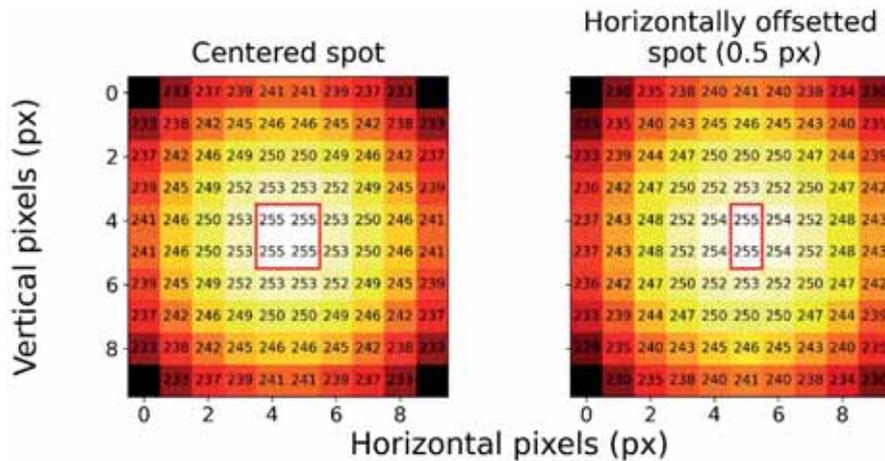


Abb. 2: Zentraler Ausschnitt der Größe 10×10 Pixel eines simulierten Lichtspots. In der linken Heatmap liegt das Spotzentrum im Ursprung (Mittelpunkt des Bildes), in der rechten ist das Zentrum um 0,5 Pixel nach rechts verschoben. Die numerischen Werte innerhalb der Kästchen stellen die Intensität der einzelnen Pixel dar. 0 bedeutet kein Licht, 255 maximale Sättigung.

2.4 Bisherige Ansätze zur Bestimmung des Spotzentrums

Wie bereits erwähnt, nutzt MAPS aktuell zwei verschiedene Algorithmen zur Bestimmung der Spotzentren. Diese sind in einer eigens entwickelten Python-Software implementiert. Der Gauß-Algorithmus nutzt das *Curve-Fitting* (Ausgleichsrechnung) mit der Gaußfunktion als Parameter. Die Funktion gibt direkt das Spotzentrum zurück. Beim Moments-Algorithmus wird das Bild zuerst binarisiert, dies erleichtert die Bestimmung der Kanten des Lichtspots im Bild. Anschließend wird die Methode der Momente angewandt, welche dann wiederum das Spotzentrum berechnet.

Die Performance beider Algorithmen wurde anhand von 1.000 simulierten Spots getestet. Das Spotzentrum wurde im Bereich 0 bis 1 Pixel, rechts vom Ursprung variiert, um ein volles Pixelspektrum abzudecken. In Abbildung 3 ist ein direkter Vergleich der beiden Algorithmen aufgezeichnet. Es ist wichtig zu vermerken, dass unter Verwendung des Moments-Algorithmus ein ausgeprägter systematischer Fehler einhergeht. Wird der Gauß-Algorithmus verwendet, kann nur ein zufälliger, wesentlich kleinerer Fehler beobachtet werden.

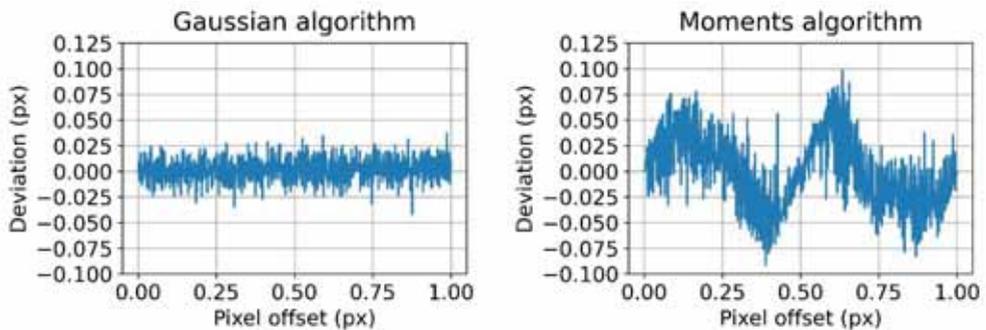


Abb. 3: Abweichung (Deviation) in Pixel zum tatsächlichen Spotzentrum unter Verwendung des Gauß-Algorithmus (links) und des Moments-Algorithmus (rechts). In den Graphen ist die Abweichung zum Sollwert über den Pixeloffset zum Ursprung aufgetragen. Das Spotzentrum wurde in der Simulation im Bereich 0 bis 1 Pixel rechts vom Ursprung variiert. Simuliert wurden 1.000 Spots mit angewandtem Gauß- und Fourier Filter.

2.5 Bestimmung des Spotzentrums mit Neuronalem Netz

Das Neuronale Netz soll anhand von Bildern der Spots trainiert werden. Dabei wird ein sogenanntes bildregressives Neuronales Netz verwendet, welches Bilder als Eingabedaten erwartet und diskrete Werte ausgibt (CHOLLET & OTHERS 2015). Für das Training sind Bilder mit bekannter Position des Spotzentrums notwendig. Bei echten Messgerätaufnahmen sind diese nicht verfügbar, sondern werden bisher durch die Algorithmen bestimmt. Die von den Algorithmen bestimmten Zentren können allerdings nicht als Label benutzt werden, da das NN sonst deren Performance nie überbieten kann. Aus diesem Grund wird auf die Simulation der Lichtspots zurückgegriffen. Sie bietet den Vorteil, dass beliebig viele Bilder mit unterschiedlichen Parameterwerten simuliert und gespeichert werden können. Die daraus entste-