

# Posenbestimmung von Schreibgeräten

Thema:

Posenbestimmung von Schreibgeräten

Art:

MA

BetreuerIn:

Florin Schwappach

BearbeiterIn:

Marco Ziegauß

ErstgutachterIn:

Christian Wolff

ZweitgutachterIn:

N.N.

Status:

abgeschlossen

Stichworte:

3D, CNN, Computer Vision

angelegt:

2018-11-26

Beginn:

2018-12-17

Anmeldung:

2018-12-17

Antrittsvortrag:

2018-12-17

Abschlussvortrag:

2019-07-01

Textlizenz:

Unbekannt

Codelizenz:

Unbekannt

## Hintergrund

Trotz fortschreitender Digitalisierung sind Stift und Papier, aufgrund ihrer einzigartigen Vorzüge, nicht aus dem beruflichen wie privaten Alltag wegzudenken (Sellen et.al, 2001). Für die Überführung handschriftlicher Notizen und Informationen in ein Computersystem ist meist jedoch mühsame Handarbeit notwendig, weshalb seit Jahren Bemühungen existieren beide Medien sinnvoller und effizienter miteinander zu vereinen (Guimbretière, 2003). Durch die Verwendung von maschinellem Lernen konnten diesbezüglich einige Fortschritte in den Bereichen der elektronischen Dokumentsegmentierung sowie der optischen Zeichenerkennung erzielt werden (Carton et. al, 2017). Andere Ansätze verfolgen das Ziel, die Position und Orientierung der verwendeten Schreibgeräte mithilfe eines Kamerasystems zu erfassen, um so tiefer gehende Informationen über Bewegungsmuster während des Schreibens zu gewinnen. Die neuesten Erkenntnisse der Objektposedetermination aus dem Bereich des maschinellen Sehens versuchen hier Defizite zu beheben und die Verwendbarkeit in Echtzeitsystemen zu erhöhen, was insbesondere für

Anwendungen der Augmented-Reality, Robotik und Human-Computer-Interaction von großer Bedeutung ist (Cao et.al, 2003).

## Zielsetzung der Arbeit

Im Rahmen dieser Arbeit sollen Systeme aus dem Bereich des maschinellen Sehens mit Convolutional Neural Networks, im Hinblick auf die Posenbestimmung von Schreibgeräten, trainiert und evaluiert werden. Dazu notwendig ist die Sammlung, Analyse und Identifikation verwendbarer Ansätze aus der Literatur, welche sich auf die oben beschriebene Anwendungsdomäne adaptieren lassen. Ziel ist es, die trainierten Systeme, mithilfe geeigneter Metriken, dahin gehend zu untersuchen, inwiefern diese zur Lösung der Problemstellung geeignet sind und mit welchen Vor- und Nachteilen sie sich voneinander Abgrenzen. Durch die steigende Leistungsfähigkeit der oft dafür verwendeten Grafikkarten ist die Forschung hinsichtlich der Anwendbarkeit in Echtzeitsystemen besonders interessant. Das Ziel ist eine Basis für neue Interaktions- und Eingabemöglichkeiten zu schaffen. Da für das Training derartiger Netzwerke stets eine existierende Datenbasis notwendig ist, ist die Entwicklung und Erstellung eines kontextbezogenen Datensatzes ebenfalls ein großer Bestandteil der Arbeit. Die Methoden zur Datenbeschaffung, die Kriterien der Datenannotation sowie das zu verwendende Format mit dem Ziel der Wiederverwendbarkeit sollen identifiziert und definiert werden.

## Konkrete Aufgaben

Konkrete Aufgaben:

- Einarbeitung in das Thema Objektposenerkennung
- Systematischer Überblick sowie Analyse existierender Ansätze hinsichtlich der Verwendbarkeit in der Problemdomäne
- Herausarbeiten der Anforderungen an den zu erstellenden Datensatz
- Erstellung eines wiederverwendbaren und kontextspezifischen Datensatzes
- Auswahl, Anpassung und Training geeigneter Deep-Learning-Netzwerke
- Identifikation geeigneter Evaluationsmetriken
- Evaluation und Vergleich der Netzwerke
- Evaluation des erstellten Datensatzes

## Erwartete Vorkenntnisse

- Erfahrungen mit Machine-Learning-Systemen und entsprechenden Frameworks
- Bildverarbeitung und maschinelles Sehen
- Gute Programmierkenntnisse
- Erfahrung mit C++, C#, Python

## Weiterführende Quellen

- Wei, S. E., Ramakrishna, V., Kanade, T., & Sheikh, Y. (2016). Convolutional pose machines. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4724-4732).
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image

recognition. arXiv:1409.1556.

- Do, T., Pham, T., Cai, M., & Reid, I.D. (2018). Real-time Monocular Object Instance 6D Pose Estimation. BMVC.
- Munich, M. E., & Perona, P. (1996, September). Visual input for pen-based computers. In Image Processing, 1996. Proceedings., International Conference on (Vol. 2, pp. 173-176). IEEE.
- Seok, J. H., Levasseur, S., Kim, K. E., & Kim, J. (2008). Tracing handwriting on paper document under video camera. ICFHR.
- Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2016). Realtime multi-person 2d pose estimation using part affinity fields. arXiv preprint arXiv:1611.08050.
- Martinez, J., Hossain, R., Romero, J., & Little, J. J. (2017, May). A simple yet effective baseline for 3d human pose estimation. In International Conference on Computer Vision (Vol. 1, No. 2, p. 5).
- Sellen, A. J., & Harper, R. H. (2003). The myth of the paperless office. MIT press.
- Guimbretière, F. (2003). Paper augmented digital documents. In: Proceedings of the 16th annual ACM symposium on User interface software and technology (pp. 51-60). ACM.
- Carton, C., Lemaitre, A., & Coüasnon, B. (2017). Eyes Wide Open: an interactive learning method for the design of rule-based systems. In: International Journal on Document Analysis and Recognition (IJ DAR), 20(2), 91-103.
- Cao, X., & Balakrishnan, R. (2003). VisionWand: interaction techniques for large displays using a passive wand tracked in 3D. In: Proceedings of the 16th annual ACM symposium on User interface software and technology (pp. 173-182). ACM.

From:

<https://wiki.mi.ur.de/> - **MI Wiki**

Permanent link:

[https://wiki.mi.ur.de/arbeiten/posenbestimmung\\_von\\_schreibgeraeten](https://wiki.mi.ur.de/arbeiten/posenbestimmung_von_schreibgeraeten)

Last update: **01.10.2019 12:02**

