

Gewährleistung der Bildkonsistenz in der Portaitfotografie: Eine technologische Lösung für effiziente und vergleichbare Aufnahmen

Willy Palme¹, Janine Breithecker¹, Olaf Rieckmann², Alexander Stolpmann¹

Zusammenfassung

Dieses Paper präsentiert das Forschungsprojekt „Intelligent Camera Unit“ (ICU) der Technischen Hochschule Wildau in Kooperation mit der Beiersdorf AG zur Verbesserung von Portraitaufnahmen in Probandenstudien für Kosmetikprodukte. Das Hauptziel des Projekt besteht darin, Ausrichtungsfehler zu minimieren und die Vergleichbarkeit von Vorher-Nachher-Bildern zu erhöhen. Dies wird durch die Verwendung eines kollaborativen Roboters und KI-gestützter Bildanalyse erreicht, um präzise Ausrichtung und Gesichtspositionen der Probanden zu gewährleisten. Das entwickelte System ermöglicht effiziente und reproduzierbare Aufnahmen aus verschiedenen Winkeln und Entfernungen und bietet eine benutzerfreundliche, web-basierte Bedienoberfläche.

Stichwörter

Künstliche Intelligenz, Kollaborative Roboter, Automatisierte Fotografie, Bildverarbeitung

1 Projektziel

Das Projekt verfolgt das Ziel, den Prozess der Portraitaufnahmen in Probandenstudien zur Erforschung von Kosmetikprodukten effizienter zu gestalten, während gleichzeitig Abweichungen in den Aufnahmen, die durch Ausrichtungsfehler verursacht werden, reduziert werden. Dies wird angestrebt, um die Vergleichbarkeit von Vorher-Nachher-Bildern zu verbessern, die oft über mehrere Wochen oder Monate hinweg entstehen können. Diese Bilder dienen der Untersuchung und Validierung der Wirkung der getesteten Produkte.

2 Konzept

Für das Grundkonzept wurde die in Bild 1 dargestellte Skizze erstellt. Dafür sieht das grundlegende Konzept vor, dass eine Fotokamera an einem Roboterarm eines kollaborativen Roboters (Cobot) befestigt wird, wie es in Bild 1 (a) dargestellt ist, damit sich diese automatisch anhand der Kopfposition und Orientierung ausrichten kann, damit der auf dem Foto abgebildete Kopf immer dieselbe Ausrichtung besitzt, wie es in Bild 1 (b) abgebildet ist.

Bei korrekter Ausrichtung wird anschließend automatisch ein Signal an die Fotokamera gesendet, die daraufhin ein Foto aufnimmt. Dabei ist es essenziell, die Kopfposition und Orientierung des Probanden akkurat und in Echtzeit bestimmen zu können. Hierfür wird ein Machine-Learning-Modell aus

¹ Technische Hochschule Wildau
² Beiersdorf AG

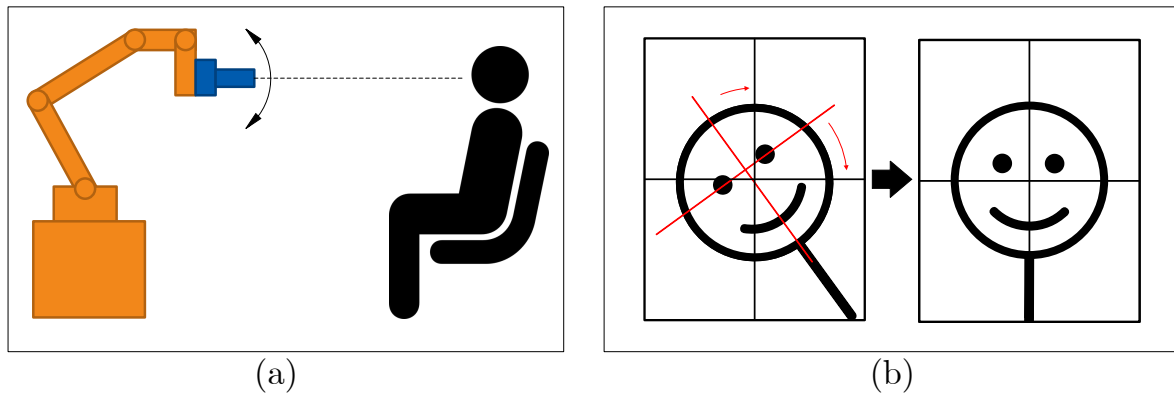


Bild 1: Skizze des Grundkonzepts

dem Open-Source Google Mediapipe Framework eingesetzt, dessen nähere Details in Kapitel 3.2 erläutert werden. Anhand der bestimmten Kopfposition und -orientierung lassen sich dann die Roboter-Zielkoordinaten für den Roboterarm, an dem sich die Fotokamera befindet, bestimmen, und es kann ein entsprechender Steuerungsbefehl an den Cobot gesendet werden.

3 Umsetzung

In diesem Kapitel wird auf die Umsetzung des Systems eingegangen. Dabei wird die grundlegende Systemstruktur sowie der finale Aufbau des ICU-Systems beschrieben.

3.1 Systemstruktur

Das System besteht, wie im Verteilungsdiagramm (Bild 2) dargestellt, auf der Hardwareseite aus einem zentralen Steuerungs-PC, an dem die restlichen Hardwarekomponenten angeschlossen werden. Dieser PC ist dafür zuständig, alle notwendige Software des Systems auszuführen und die einzelnen Hardwarekomponenten anzusteuern. Über das Ethernet ist ein UR5 Cobot, der Firma Universal Robots, mit dem PC verbunden. Dieser ist dafür zuständig, die Fotokamera exakt nach den Vorgaben des Softwaresystems vor dem Probanden zu platzieren.

Des Weiteren ist eine Intel Realsense D415 Tiefenbildkamera über USB mit dem Steuerungs-PC verbunden. Diese wird vom Softwaresystem dafür genutzt, die aktuelle Position und Orientierung des Probanden im Verhältnis zur gegenwärtigen Position des Roboterarms zu bestimmen. Die letzte Hauptkomponente ist die über USB verbundene Systemkamera aus der „R“-Serie der Firma Canon. Diese wird dazu genutzt, die gewünschten hochauflösenden Bilder des Probanden aufzunehmen.

Optional kann auch noch eine beliebige Anzahl an Client-Geräten zur Steuerung des Systems über Ethernet oder auch WLAN verbunden werden. Auf dem Hauptsteuerungs-PC wird das für dieses Projekt entwickelte Softwarepaket ausgeführt. Die Software wurde in Python entwickelt und nutzt derzeit die Python Version 3.10. Softwareseitig wird die Canon Kamera mithilfe der Canon EOS Digital SDK [1] angesprochen und der UR5 Cobot wird mithilfe des UR-RTDE Frameworks [2] von SDU Robotics gesteuert. Die Software stellt einen UI-Webserver bereit, der es einer beliebigen Anzahl von Client-Geräten ermöglicht, das System zu steuern.

Die Software, die für das Projekt entwickelt wurde, ist ein komplexes System, das insgesamt aus über 8000 Zeilen Programmcode besteht. Dabei wurde darauf geachtet, dass das System so weit wie möglich modular aufgebaut ist, um gegebenenfalls zukünftig gewünschte oder notwendige Anpassungen schnell vornehmen zu können. Die Software nutzt zudem Multi-Threading, um das System performant und gleichzeitig reaktionsfreudig bedienbar zu halten. Das Nutzen von Multi-Threading hat die Komplexität des Systems wesentlich erhöht, ist aber ein notwendiger Bestandteil, um das System in der aktuellen Form umzusetzen.

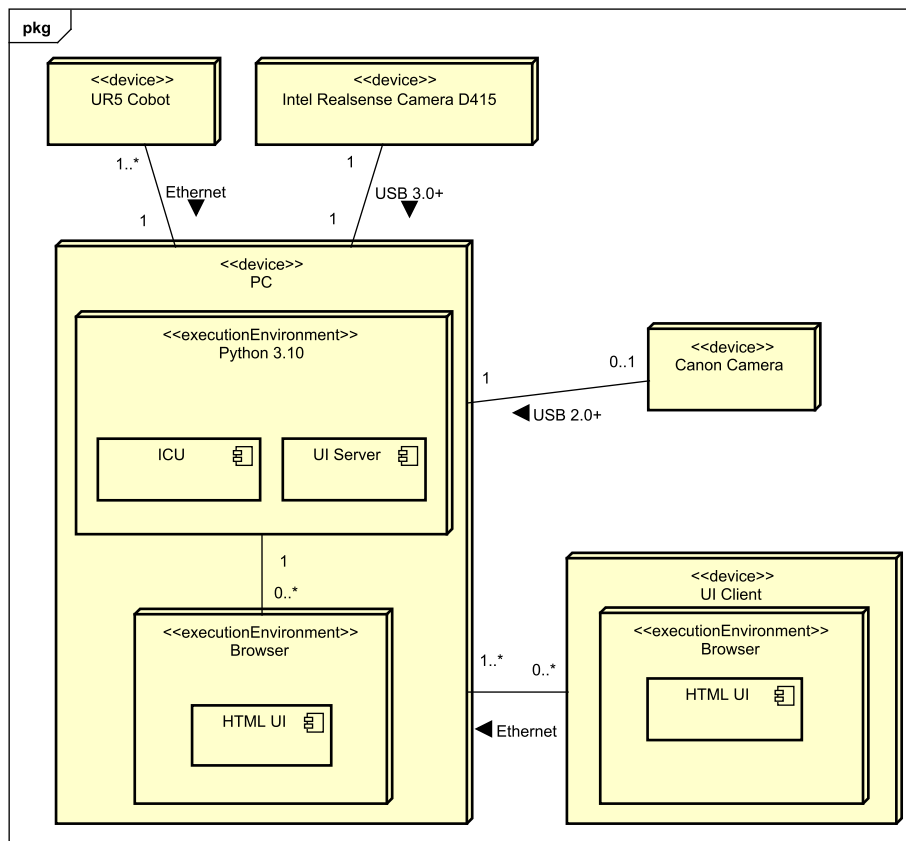


Bild 2: Verteilungsdiagramm

3.2 Finaler Aufbau

Zur Umsetzung des Projekts wurde eine professionelle spiegellose Systemkamera (Canon EOS RP) an einen kollaborativen Roboter (Universal Robots UR5) montiert, um eine präzise Ausrichtung der Fotokamera im Raum sicherzustellen. Damit vergleichbare Portraitaufnahmen erreicht werden, ist es nicht nur notwendig, die exakte Positionierung der Kamera zu gewährleisten, sondern auch entscheidend, die Gesichtsposition und -ausrichtung der Probanden zur Kamera zu bestimmen. Der finale Aufbau ist in Bild 3 zu sehen.



Bild 3: Finaler Aufbau

Die Gesichtspositionierung und -orientierung wird aus dem Zusammenspiel eines „Face Landmarkers“ aus dem Google Mediapipe Framework [3] und einer Intel Realsense Tiefenbildkamera, die über die Intel Realsense SDK [4] angesprochen wird, bestimmt. Der „Face Landmarkers“ besteht dabei aus drei Maschine-Learning-Modellen: dem „BlazeFace-“ [5, 6], für die Gesichtserkennung, dem „FaceMesh-“ [7], für die Orientierungspunkt-Erkennung und dem „BlendShape-Model“ [8, 9], für das Erkennen von verschiedenen Gesichtsausdrücken. Dies ermöglicht die präzise Positionierung der Fotokamera, sodass die Aufnahmen immer aus dem gleichen Winkel und der gleichen Entfernung zum Gesicht des Probanden aufgenommen werden.

Der „Face Landmarker“ erlaubt auch die Auswertung der Gesichtsausdrücke der Probanden während der Ausrichtung der Kamera sowie in den aufgenommenen Fotos. Auf diese Weise wird sichergestellt, dass die Gesichtsausdrücke auf den Probandenbildern den gewünschten Kriterien entsprechen und zwischen den Vorher-Nachher-Aufnahmen (Bild 4) konsistent sind. Diese Kriterien können beispielsweise die geschlossenen Augen und einen neutralen Gesichtsausdruck des Probanden umfassen.



Bild 4: Exemplarische Vergleichsbilder

Das ICU-System ermöglicht es effizient, eine Reihe von Fotos reproduzierbar aus verschiedenen konfigurierbaren Aufnahmewinkeln und Entfernungen aufzunehmen.

Um das System einfach und effizient bedienen zu können, wurde eine web-basierte Bedienoberfläche (User Interface – UI) (Bild 5) entwickelt, die es ermöglicht, das System lokal oder remote über jeden beliebigen Webbrowser zu bedienen. Das UI bietet eine Übersicht über den Status der verschiedenen Teilsysteme und eine Live-Ansicht des Kamerastreams der Tiefenbildkamera, mithilfe dessen die Ausrichtung des Systems vorgenommen wird. In dem Kamerastream wird zudem gezeigt, ob der Face-Landmarker ein Gesicht erkannt hat und ob die Gesichtsausdrücke (Augen offen/geschlossen, Lächeln ja/nein) des Probanden den in der Konfigurationsdatei eingegebenen Vorgaben entsprechen.

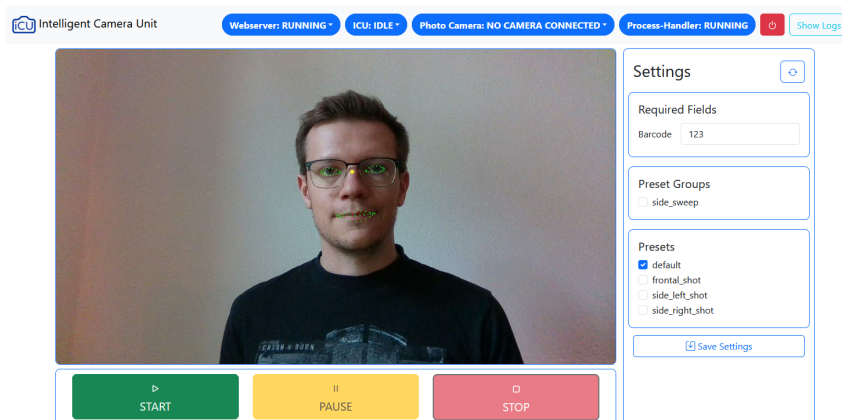


Bild 5: Webbasiertes Frontend

Wenn eine Aufnahme von dem Studiendurchführenden gestartet wird, schaut das System zunächst, ob der Proband innerhalb des definierten Aufnahmebereichs sitzt und grob in Richtung der Fotokamera

schaut. Dies ist empfehlenswert, da das System sonst Steuerungsbefehle an den Cobot senden könnte, die außerhalb des Arbeitsbereiches des Cobots liegen. Ist die Grundausrichtung des Probanden erreicht, bestimmt das System die Position und Orientierung des Kopfes des Probanden und berechnet die Zielkoordinaten für den Roboterarm, um ein exakt ausgerichtetes Foto zu bestimmen. Befindet sich der Roboterarm nicht in den optimalen Koordinaten, werden die Zielkoordinaten an den Cobot übertragen, der sich dann dorthin bewegt. Dieser Schritt wird so oft wiederholt, bis die berechneten Zielkoordinaten den aktuellen Koordinaten des Roboterarms entsprechen und das System somit genau ausgerichtet ist. Wenn das System ausgerichtet ist, wird ein Signal an die Fotokamera geschickt, welche dann ein Foto des Probanden macht. Das aufgenommene Foto wird an das System übertragen und zunächst überprüft, ob die Gesichtsausdrücke und die Orientierung im finalen Bild den Vorgaben in der Konfigurationsdatei entsprechen. Sind die Vorgaben nicht erfüllt, wird die Aufnahme wiederholt. Bei Erfüllung der Vorgaben wird das finale Bild dem Bediener zur Bestätigung angezeigt, sodass dieser überprüfen kann, ob das Foto korrekt aufgenommen wurde oder es wiederholt werden muss. Dabei ist es für den Bediener auch möglich, sich ein Referenzbild Seite an Seite mit dem gerade aufgenommenen Foto anzeigen zu lassen, um eventuelle Abweichungen zu erkennen (siehe Bild 6). Wird das Foto bestätigt, speichert das System das Foto sowie die Augmentierungsdaten der Gesichtsausdrucks auf dem System ab, und die Aufnahme ist abgeschlossen.

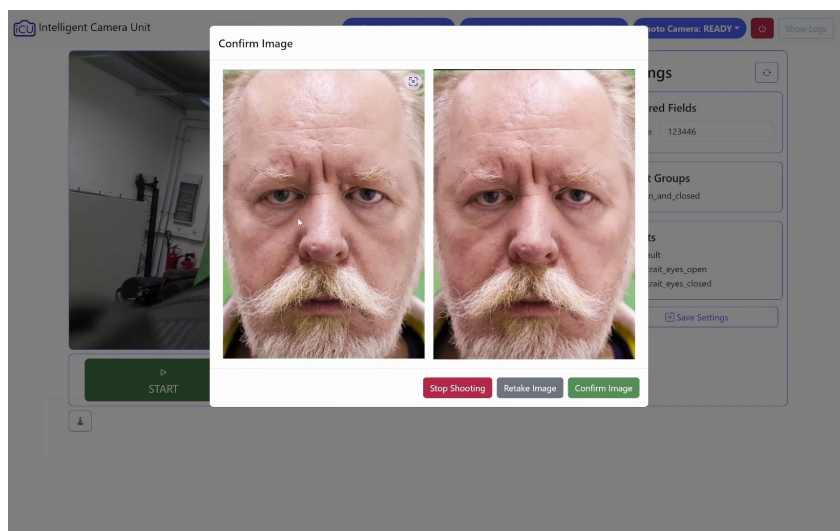


Bild 6: Referenzansicht der Vorher-Nachher-Fotografien

4 Fazit

Zusammenfassend präsentiert dieses Paper das Forschungsprojekt „Intelligent Camera Unit“ (ICU) der Technischen Hochschule Wildau in Zusammenarbeit mit der Beiersdorf AG. Das Hauptziel des Projekts besteht darin, Portraitaufnahmen in Probandenstudien für Kosmetikprodukte zu verbessern, indem Ausrichtungsfehler minimiert und die Vergleichbarkeit von Vorher-Nachher-Bildern erhöht werden. Dies wird durch den Einsatz eines kollaborativen Roboters und KI-gestützter Bildanalyse erreicht, um präzise Ausrichtung und Gesichtspeditionen der Probanden sicherzustellen.

Die Kombination aus einem kollaborativen Roboter und KI-basierter Bildanalyse ermöglicht präzise Ausrichtung und Gesichtspeditionen. Das System gewährleistet effiziente und reproduzierbare Aufnahmen aus verschiedenen Winkeln und Entfernungen.

Insgesamt präsentiert das Paper eine innovative Lösung für die Optimierung von Portraitaufnahmen in Probandenstudien, indem es modernste Technologien wie kollaborative Roboter, KI und Bildverarbeitung integriert. Der vorgestellte Ansatz bietet nicht nur effiziente und reproduzierbare Ergebnisse, sondern auch eine benutzerfreundliche webbasierte Steuerungsschnittstelle für Anwender.

5 Literatur

- [1] M. Hawkins, “Understanding Canon’s EOS Digital SDK.” [Online]. Available: <https://en.canon-cna.com/pro/stories/eos-digital-sdk-explained/>
- [2] A. P. Lindvig, “SDU robotics - ur-rtde - GITLAB,” 2023. [Online]. Available: https://gitlab.com/sdurobotics/ur_rtde
- [3] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann, “MediaPipe: A Framework for Perceiving and Processing Reality,” in *Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR) 2019*, 2019. [Online]. Available: https://mixedreality.cs.cornell.edu/s/NewTitle_May1_MediaPipe_CVPR_CV4ARVR_Workshop_2019.pdf
- [4] “Intel RealSense SDK 2.0,” 2023. [Online]. Available: <https://www.intelrealsense.com/sdk-2/>
- [5] V. Bazarevsky, Y. Kartynnik, and A. Ablavatski, “MediaPipe BlazeFace Model Card,” 2021. [Online]. Available: [https://storage.googleapis.com/mediapipe-assets/MediaPipe%20BlazeFace%20Model%20Card%20\(Short%20Range\).pdf](https://storage.googleapis.com/mediapipe-assets/MediaPipe%20BlazeFace%20Model%20Card%20(Short%20Range).pdf)
- [6] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, “BlazePose: On-device Real-time Body Pose tracking,” 2020.
- [7] I. Grishchenko and G. Yan, “MediaPipe FaceMesh Model Card,” 2022. [Online]. Available: <https://storage.googleapis.com/mediapipe-assets/Model%20Card%20MediaPipe%20Face%20Mesh%20V2.pdf>
- [8] I. Grishchenko, G. Yan, A. Zanfir, and E. G. Bazavan, “MediaPipe Blendshape V2 Model Card,” 2022. [Online]. Available: <https://storage.googleapis.com/mediapipe-assets/Model%20Card%20Blendshape%20V2.pdf>
- [9] I. Grishchenko, G. Yan, E. G. Bazavan, A. Zanfir, N. Chinaev, K. Raveendran, M. Grundmann, and C. Sminchisescu, “Blendshapes GHUM: Real-time Monocular Facial Blendshape Prediction,” 2023.