

Photogrammetrische Bestimmung von Kopfbewegungen mittels Facial Landmarks – proof of concept

Dr.-Ing. Mirko Riedel (deveritec), Prof. Dr.-Ing. Danilo Schneider (HTW Dresden),
Dr.-Ing. Anne Bienert (TU Dresden)

deveritec GmbH
Bamberger Straße 1, 01187 Dresden
eMail: mirko.riedel@deveritec.com
URL: <http://www.deveritec.de>

Zusammenfassung.

Im Rahmen der Krebsbehandlung im Kopfbereich ist die Strahlentherapie ein etabliertes Verfahren. Um während einer Therapiesitzung eine Bewegung des Kopfes zu verhindern, werden derzeit mechanische Fixierungen (Kopfmasken) verwendet. Diese verursachen eine starke mechanische und psychische Belastung des Patienten. Das Ziel des Forschungsprojektes „Increasing the precision of patient positioning in radiotherapy using positioning aids“ (imPRESSing) ist deshalb eine Ablösung der mechanischen Fixierung durch eine messtechnische Überwachung der Kopfposition und -neigung. Der Schwerpunkt dieses Beitrags liegt dabei auf dem proof of concept der optischen Bewegungsdetektion. Dabei wird auf den Aufbau und Funktionsweise, die Bildmessung der „Facial Landmarks“ sowie die Systemkalibrierung eingegangen. Weiterhin werden erste experimentelle Ergebnisse der Kalibrierung, der Bildmessung und der Bestimmung der 3D-Koordinaten vorgestellt. Diese zeigen, dass die Bildmessung prinzipiell funktioniert und die photogrammetrische Kopfposenbestimmung mit einer ausreichenden Genauigkeit (im Bereich weniger Millimeter) bestimmt werden können. Gleichzeitig zeigt sich auch weiteres Verbesserungspotenzial. Auf Basis der ersten Ergebnisse wird auch auf die erreichbare Genauigkeit und Unsicherheit sowie auf den Einfluss von äußeren Randbedingungen eingegangen.

1 Einleitung

Das Ziel der Strahlentherapie ist ein Abtöten des Tumorgewebes durch eine lokal konzentrierte, hochenergetische Bestrahlung. Eine Therapie umfasst bis zu 30 einzelne Therapiesitzungen in kurzer zeitlicher Folge mit einer Bestrahlungsdauer von jeweils weniger als 10 Minuten. In einem vorgelagerten Planungsschritt wird die Bestrahlungsleistung und -richtung durch eine Anpassung der Bewegung der Strahleneinheit um den Patienten herum optimiert, sodass der Strahlungseintrag im angrenzenden Gewebe minimiert wird (**Abb. 1**, rechts oben).

Das Bestrahlungssystem mit dem Strahlaustritt (**Abb. 1**, links) ist in einem nach außen hin strahlungsschützendem Raum mit einem ortsfesten Bezugskoordinatensystem installiert. Am Bestrahlungssystem befindet sich ein fest installierter und mitbewegter Computertomograph (CT). Um den Weg der Strahlung zu variieren kann das Gesamtsystem dabei um die Y-Achse (grüne Achse) rotieren. Weiterhin kann die

Patientenliege in ihrer Position und Neigung eine Feinkorrektur mit eingeschränktem Bewegungsbereich sowie die Liegenbasis eine Rotation um ihre Z-Achse (blaue Achse) in einem großen Stellbereich (max. $\pm 90^\circ$) realisieren. Für eine Bestrahlung wird der Patient auf der Patientenliege so positioniert, dass der Tumor sich im Isozentrum der Bestrahlungseinrichtung befindet. Mittels des Kontroll-CTs werden Position und Größe des Tumors gegenüber der Diagnose kontrolliert. Gegebenenfalls erfolgt eine Feinpositionierung der Liege (6DoF Bewegungseinrichtung) um kleinere Abweichungen der Patientenpositionierung auszugleichen.

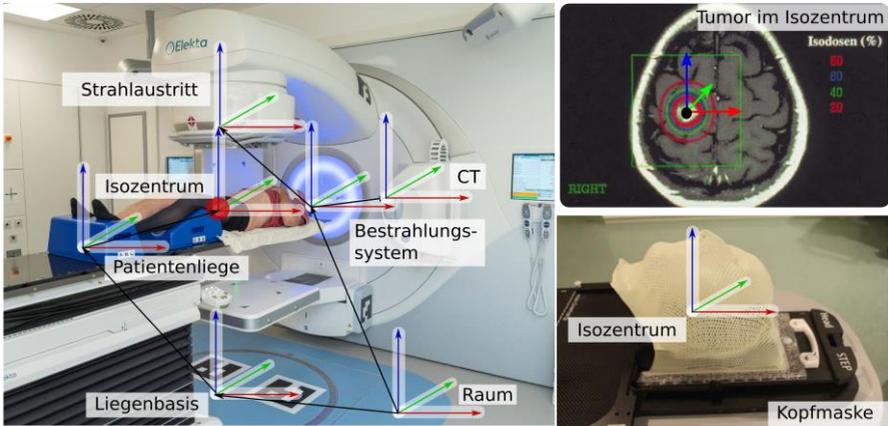


Abb. 1: links: Komponenten eines Systems zur Strahlentherapie, rechts unten: Kopfmaske zur medizinischen Kopffixierung, rechts oben: Ausrichtung des Tumors im Isozentrum der Anlage mit simulierter Verteilung der Strahlungsleistung (vgl. 1 und 2).

Eine essentielle Voraussetzung ist die Vermeidung einer Bewegung des Kopfes durch eine mechanische Fixierung (**Abb. 1**, rechts unten) während der Bestrahlung. Diese belastet den Patienten jedoch körperlich und psychisch stark und verursacht für das medizinische Personal einen erhöhten Arbeitsaufwand zur initialen Erstellung sowie beim Anbringen in jeder Therapiesitzung.

2 Projekt „imPRESSing“

Um diese Defizite zu beseitigen, wurde das Forschungsprojekt imPRESSing initiiert. Das Hauptziel ist es, die mechanische Fixierung durch eine stabile Lagerung in Kombination mit einer messtechnischen Überwachung der Kopfposition zu ersetzen. Um ein maximal hohes Sicherheitsniveau für die messtechnische Überwachung der Kopfposition zu erreichen, ist im Projekt die Kombination von drei unabhängigen Lösungsansätzen vorgesehen. Diese sind, mittels einem am Patienten angebrachten Tag die Funkortung auf Basis von UWB-RTLS (Ultra- Wideband Real-time locating system) in Kombination mit Beschleunigungsaufnehmern und einer IMU (Inertial Measurement Unit) sowie einem optischen Tracking, d.h. der kontinuierlichen photogrammetrischen Posenbestimmung (**Abb. 2**). Aus Anwendungssicht hat das optische Tracking im Vergleich zu den alternativen Ansätzen einen erheblichen

Vorteil, da es theoretisch ohne Marken mit einer ausreichend hohen Genauigkeit und Zuverlässigkeit arbeiten könnte. Um dies zu untersuchen, wurde ein „proof of concept“ des optischen Systems initiiert, der den Schwerpunkt dieses Beitrags bildet.

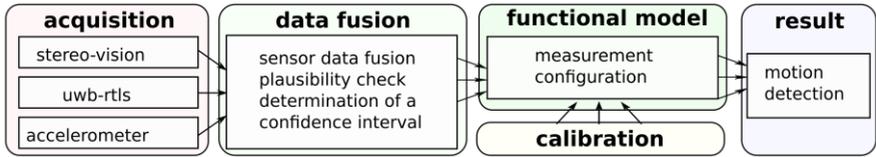


Abb. 2: Multisensor-Lösungsansatz „imPRESSing“ mit stereo-vision, UWB-RTLS sowie einer IMU.

3 Systemaufbau und Funktionsweise

Nicht nur aus datenschutzrechtlichen Gründen wurde für die initialen Versuche ein Kopf-Dummy in realistischer Größe (**Abb. 3**) verwendet. Aus Vorversuchen ist bekannt, dass die Erkennungsrate durch die fehlende Hauttextur etwas geringer als bei einem Menschen, aber prinzipiell ausreichend hoch ist.



Abb. 3: Blickfeld und Kopfbereich im Bild einer der untersuchten Kameraanordnungen sowie die verwendeten Kalibrierkörper.

Für die Bildaufnahme werden einfache, nicht synchronisierte Kameras mit einer nominellen Auflösung von 2560 x 1440 Pixel verwendet. Damit lässt sich ein gut überlappendes Blickfeld der Augen- und Kinnpartie des Kopfes sowie der Patientenliege mit dem Hintergrundpattern (aus ArUco-Markern) erfassen (**Abb. 3**, links). Der Roboter dient zur Verifikation der Messergebnisse der Bewegungsdetektion (**Abb. 3**, mittig). Die Wiederholgenauigkeit beträgt laut Hersteller ca. 0.1 mm im Raum, sodass auch kleine Bewegungen und Rotationen des Kopfes mit ausreichender Präzision definiert simuliert werden können.

4 Systemkalibrierung

Für die Systemkalibrierung wurden die Standard OpenCV-Kalibrierkörper Chessboard und ArUco (Abb. 3, rechts) sowie ein robuster und beweglicher, im Vorfeld kalibrierter Referenzkörper in Anlehnung an VDI 2634 [3] mit 14 Referenzstrecken (Abb. 4) verwendet. Die Referenzkörper entsprechen dabei der Größe des Kopf-Dummys und sind größer als die zu überwachenden Gesichtspartien.



Abb. 4 Referenzkörper (Abmessungen: 270 mm x 250 mm x 130 mm) mit Kontrollstrecken zur Systemkalibrierung im Größenvergleich zum verwendeten Kopf-Dummy.

Bestimmt werden dabei die inneren und relativen Orientierungen der Kameras zueinander, sowie die relative Orientierung des Kamerasystems zur Patiententiege über die darauf angebrachten Marker des Hintergrundpatterns (vgl. Abb. 3). Zur Kalibrierung wird dabei der Probekörper manuell im Zielmessvolumen bewegt. Um eine gute und gleichmäßige Abdeckung der Bilder aller Kameras zu erreichen, werden zwischen 15 und 30 Bildsätze aller Kameras mit unterschiedlichen Aufnahmepositionen des Referenzkörpers verwendet. Zielkriterien sind dabei eine möglichst gute Abdeckung des Bildes jeder Kamera sowie des Objektraumes in dem der Kopf sich später befindet.

Tabelle 1: Innere Orientierung unterschiedlicher Kameras und Kalibrierkörper.

		Depstech 1280x720 Pixel				General Webcam 2560x1440 Pixel	
		Chessboard - eben		ArUco - eben		ArUco - räumlich	
		Wert	σ	Wert	σ	Wert	σ
fx	Pixel	721.66	0.31	715.97	0.93	1906.63	0.79
fy	Pixel	721.71	0.35	715.02	1.02	1906.28	0.82
cx	Pixel	641.14	0.43	624.99	1.38	1330.43	1.37
cy	Pixel	351.07	0.53	361.67	1.49	519.58	1.19
k1		0.1795	0.0018	0.1609	0.0053	-0.3711	0.0021
k2		-0.2528	0.0048	-0.2219	0.0132	0.1435	0.0049
p1		0.0004	0.0003	-0.0009	0.0007	-0.0010	0.0002
p2		-0.0007	0.0002	-0.0003	0.0007	0.0007	0.0001
k3		0.1252	0.0041	0.1006	0.0104	-0.0284	0.0033

Interessant ist dabei, dass mit einem ebenen Chessboard für Brennweite (fx, fy) und Hauptpunkt (cx, cy) die Unsicherheiten mit einem räumlichen ArUcoboards, das mit einer Unsicherheit der Objektkoordinaten von ca. 0.01 mm kalibriert wurde, vergleichbar sind (Tabelle 1). Zu erwarten wäre ein deutlicher Abfall, da das Chessboard eben ist und damit eine schlechtere Bestimmbarkeit der Parameter erlauben sollte. Deutliche Einbußen der Bestimmungsgüte der inneren Orientierung sind bei der Verwendung eines ebenen ArUcoboards zu verzeichnen, der Faktor der Unsicherheiten beträgt ca. 3, was so nicht erwartet wurde. Die äußere bzw. relative Orientierung der Kameras in Tabelle 2 zeigt annähernd das gleiche Bild. Der Unterschied zwischen dem Chessboard und dem räumlichen ArUcoboards ist vernachlässigbar klein, wenn man die unterschiedlichen Auflösungen der Kalibrierbilder berücksichtigt. Deutlich schlechter fällt auch hier die Kalibrierung mit dem ebenen ArUcoboards aus.

Tabelle 2: Äußere Orientierung unterschiedlicher Kameras und Kalibrierkörper.

		Depstech 1280x720 Pixel				General Webcam 2560x1440 Pixel	
		Chessboard - eben		ArUco - eben		ArUco - räumlich	
		Wert	σ	Wert	σ	Wert	σ
X	mm	244.71	0.46	243.63	1.10	-524.03	0.25
Y	mm	-309.60	0.39	-308.11	1.13	-405.69	0.23
Z	mm	486.71	0.37	483.88	0.94	314.92	0.25
rx	°	-157.06	0.05	-157.18	0.16	-100.21	0.04
ry	°	23.18	0.05	22.82	0.10	-1.31	0.01
rz	°	-4.73	0.02	-4.85	0.06	-81.13	0.05

Insgesamt liefert die Systemkalibrierung mit dem Chessboard und dem räumlichen ArUcoboards für diesen Anwendungsfall ausreichend gute Ergebnisse hinsichtlich der inneren und relativen Orientierung der Kameras. Zur Ergebniseinordnung muss beachtet werden, dass für den „proof of concept“ die Bildmessung der ArUco-Targets von OpenCV sowie einfache Webcams verwendet wurden. Die Subpixel-Genauigkeit ist in den Kalibrierdaten jedoch nicht erkennbar. Es ist davon auszugehen, dass in der Wahl geeigneter Marker und Bildmessenger noch ein hohes Verbesserungspotential (Faktor 5 ... 10) liegt.

Zusammengefasst wird für den „proof of concept“ von einer prinzipiellen Kalibrierbarkeit des Gesamtsystems mit erheblichem Verbesserungspotenzial ausgegangen. Insgesamt wird für das Testsystem von einem Genauigkeitspotential im Objektraum mit ca. ± 2 mm ausgegangen.

5 Bildmessung facial landmarks

Für den „proof of concept“ der facial landmarks wurde auf vorgefertigte Bibliotheken zurückgegriffen, die sich hinsichtlich der Erkennungsgüte sowie der Anzahl der Markerpunkten im Gesicht unterscheiden (**Abb. 5**, links: dlib [4], OpenCV [5], mtcnn [6] und mediapipe [7]). Prinzipiell liefern alle Operatoren bei Frontalaufnahmen stabile Ergebnisse, fallen bei seitlichen Aufnahmen jedoch unterschiedlich stark ab. Als robusteste vorgefertigte Bibliothek hat sich dabei dlib erwiesen. Es ist davon

auszugehen, dass mittels eines erweiterten Lerndatensatzes mit mehr seitlichen Aufnahmen die Ergebnisse signifikant verbessert werden können. Es muss weiterhin geprüft werden, ob und wie sich weitere sowie speziell angepasste landmarks (z.B. der Ohren) integrieren lassen.

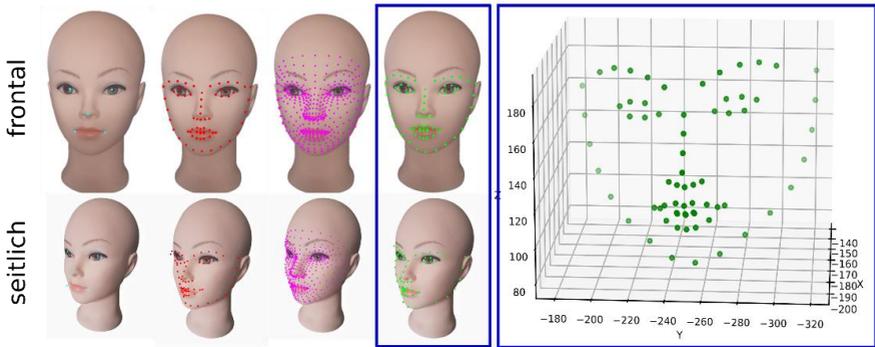


Abb. 5: links: 2D-Bildpunkte von „facial landmarks“ am Puppenkopf sowie rechts: Überführung in den Objektraum mittels räumlichem Vorwärtsschnitt.

Aus den Landmarks aller Kameras können anschließend mittels räumlichem Vorwärtsschnitt die 3D-Koordinaten der landmarks bestimmt werden (**Abb. 4**, rechts). Diese Koordinaten müssen bezüglich ihrer Lage zum Schädel als ortsfest (z.B. Nasenspitze) und beweglich (z.B. Mundpartie) unterschieden werden. Aus den ortsfesten Koordinaten kann die Lage und Orientierung des Kopfes bestimmt werden, alle anderen können zur Überwachung des allgemeinen Zustandes des Patienten verwendet werden.

5 Vorwärtsschnitt auf Basis der facial landmarks

Zur Bestimmung der facial landmarks wurden verschiedene frei verfügbare Bibliotheken getestet (dlib, OpenCV, mtcnn, mediapipe, vgl. **Abb. 5**). Alle liefern gelabelte Bildkoordinaten der Markerpunkte zurück, die mittels Vorwärtsschnitt in den 3D-Raum überführt werden können. Unterschiede der verwendeten Operatoren ergeben sich dabei hinsichtlich der Anzahl der Markerpunkte, Rechenzeit, Stabilität und Zuverlässigkeit. Beispielhaft werden hier nur die Ergebnisse des OpenCV-Operators dargestellt, da er die mit der Verwendung von facial landmarks verbundenen Herausforderungen am besten repräsentiert.

Tabelle 3: Standardabweichungen der Objektkoordinaten.

Koordinaten	XYZ	X	Y	Z
RMS	5.65 mm	4.45 mm	2.47 mm	2.45 mm
σ_{\max}	6.47 mm	5.26 mm	2.70 mm	2.64 mm

Betrachtet man die RMS-Werte aller 3D-Koordinaten der Markerpunkte im 3D-Raum

zeigen sich Standardabweichungen von ca. 6.5 mm (Tabelle 3) Dies erscheint auf den ersten Blick jedoch nur bedingt zu den Ergebnissen der Systemkalibrierung zu passen, die eine Längenmessabweichung von unter 2 mm für gut signalisierte Marken erwarten lässt. Besonders hervorstechend ist dabei die X-Koordinate (normal zur Bildebene), die relativ zu den Y- und Z-Komponenten (innerhalb der Bildebene) eine ca. doppelt so hohe Standardabweichung aufweist. Dieses Verhalten entsteht durch die Einbeziehung einer oder mehrerer fehlerhafter Bildmessungen in den räumlichen Vorwärtsschnitt.

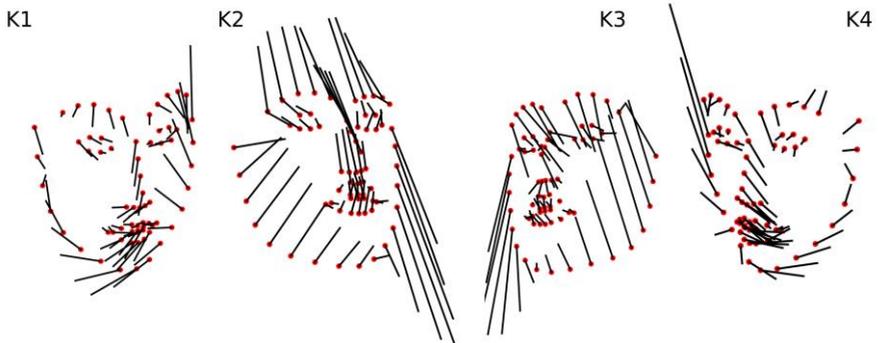


Abb. 6: RMS und Residuen der Rückprojektion von Kamera 1 bis 4 (K1-K4).

Betrachtet man die Residuen der Rückprojektion (**Abb. 6**), zeigen sich Hinweise auf die Ursachen. Dabei zeigen vor allem die Kinnpartien deutliche systematische Abweichungen. Vergleicht man diese mit den Blickfeldern der Kameras (**Abb. 3**), zeigt sich, dass die Kinnpartie in zwei von vier Kameras nicht sichtbar ist. Jedoch schätzen die facial landmarks die Kinnpartie über den Helligkeitsgradienten trotzdem mit (**Abb. 5**, links), was zu einem sehr großen Fehler führen kann.

Generell sind die Erkennungsrate und -güte für die Applikation jedoch prinzipiell tauglich, ein spezielles Anlernen ist jedoch zwingend notwendig. Dabei können auch weitere tendenziell unbewegte Features, wie beispielsweise die Ohren mit einbezogen werden. Damit können durch die besseren seitlichen Sichtbedingungen einerseits mehr Punkte einbezogen werden, was die Kontrollierbarkeit steigert als auch andererseits die Transformation auf den eigentlichen Zielpunkt, den Tumor verbessert werden.

6 Zeitliches Verhalten der facial landmarks

Ein zweiter Aspekt des „proof of concept“ ist die Abschätzung der Robustheit der Erkennung. Dazu wurden für die gezeigte statische Messkonfiguration mit einer Kamera (ähnlich **Abb. 3**) über einen längeren Zeitraum Messungen aufgenommen und die Bildkoordinaten der landmarks ausgewertet. In **Abb. 7** sind für jeweils eine Referenzmarke im Hintergrund, Nasenspitze sowie linkes und rechtes Innenaug der zeitliche Verlauf der Bildkoordinaten dargestellt.

Es zeigt sich, dass die Position der ArUco-Marke im Bereich von deutlich unter

1 Pixel über die Versuchsdauer konstant blieb. Somit kann innerhalb dieses Rahmens von einem mechanisch unbewegten Aufbau ausgegangen werden. Im Bild ergeben sich deutliche Änderungen der gemessenen Bildkoordinaten von bis zu 10 Pixeln für die X-Richtung der Nasenspitze. In Y-Richtung fallen sie mit ca. 5 Pixeln etwas geringer, jedoch immer noch deutlich höher als der ArUco-Marker aus. Überträgt man diese Bildkoordinaten mit einer abgeschätzten Ortsauflösung von ca. 0.4 mm/Pixel in den Objektraum, ergibt sich eine Schwankung des Markers für die Nasenspitze von ca. 4 mm in X-Richtung. Die Marker der Augenpunkte zeigen in X- und Y-Richtung mit ca. 2 mm im Objektraum eine etwas geringere Schwankung.

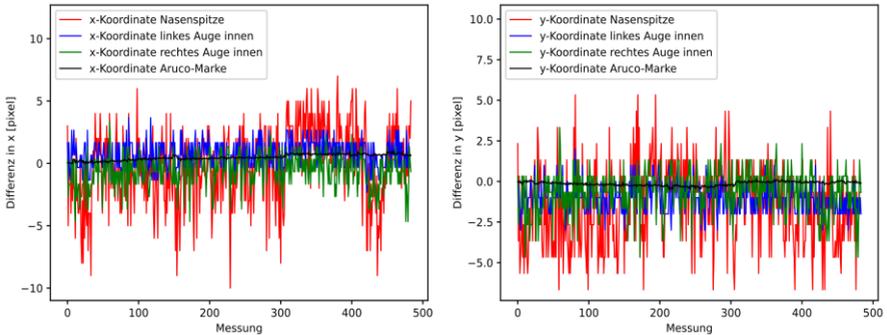


Abb. 7: Zeitliches Verhalten der Bildmessung am Beispiel von 4 ausgewählten Markern (Nasenspitze, Auge innen rechts und links, ArUco-Marke).

Die Ursachen für die relativ großen Abweichungen sind dabei noch nicht vollständig geklärt. Möglicherweise ist eine Kombination aus leicht wechselnder Beleuchtungssituation und dem oben angeführten Schätzen nicht sichtbarer Konturpunkte verantwortlich. Es ist daher davon auszugehen, dass ein besseres Anlernen der Detektion der facial landmarks die Abweichungen deutlich senken kann.

9 Vorbereitung für den klinischen Test

Aktueller Projektstand ist die Integration der Bildmessoperatoren in eine Anwendung, die im klinischen Umfeld getestet werden kann (**Abb. 8**). Darin sind links die live-Messungen mit ArUco- und facial landmarks-Messung dargestellt. Damit ist eine Kontrolle der Randbedingungen der Messung wie beispielsweise der Helligkeit, möglich.

Mittig sind die Objektkoordinaten im 3D-Raum dargestellt. Diese zeigen über die Farbcodierung die Unsicherheit der zeitlich aktuellen Bestimmung an. Hier ist zu erkennen, dass die Erkennungsgüte nur im Bereich der Nasenspitze ausreichen hoch (grün) bei der Verwendung der Mediapipe-Messung ist. Einfluss auf die Erkennungsgüte hat dabei auch der aktuelle Kalibrierzustand des Systems. Dieser wird über die Bestimmung der räumlichen Kameraposition verifiziert. Damit kann zu jeder Einzelmessung die Integrität des Systems festgestellt werden.

Der rechte Bereich enthält, vorausgesetzt das ausreichend viele hinreichend gute Punkte detektiert werden konnten, die aktuelle und die Abweichung der Kopf-Pose

gegenüber dem Messungsstart. Dabei fließen alle während der Messung erhobenen Kontrollgrößen und Unsicherheiten in die Posenbestimmung mit ein, um die Verlässlichkeit der Messung sicherstellen zu können. Der Start der Messung wird dabei durch den Bediener gesetzt, wenn der Patient so ausgerichtet ist, dass sich der Tumor im Isozentrum der Anlage befindet. In einem späteren Schritt kann durch eine Integration des Systems in die Gesamtanlage der Abgleich auch automatisiert erfolgen.

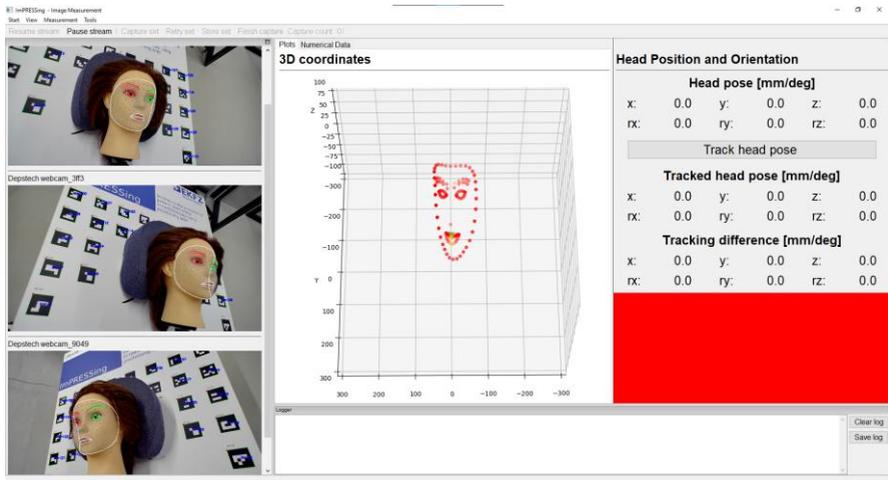


Abb. 8: Screenshot des aktuellen Ausbaustandes des Anwendungsprototyp für den Einsatz im klinischen Bereich.

Wenn die ersten Test erfolgreich verlaufen und ggf. auftretende äußere Einflüsse auf die Bildmessung beseitigt wurden, wird nur noch der rechte Bereich (**Abb. 8**) der Anwendung für den Bediener zu sehen sein.

7 Diskussion

Die Messkonfiguration ist grundsätzlich für die Messaufgabe geeignet und ist prinzipiell auch kalibrierfähig. Die Konfiguration bietet ein umfassendes Blickfeld auf den Kopf des Patienten. Einzelne Kopfparten, u.a. das Kinn, werden nicht in allen Blickwinkeln eindeutig erkannt. Die Güte der Kalibrierung ist momentan durch die Güte der eingesetzten Kameramodule sowie die Bildmessung im Subpixelbereich für ArUco-Marker der eingesetzten OpenCV-Bibliothek beschränkt. Beide Ansatzpunkte bieten jedoch Raum für eine deutliche Verbesserung der Kalibrierungsgüte.

Die Detektion von Kopfbewegungen mittels facial landmarks ist prinzipiell möglich. Bei Frontalaufnahmen (vgl. Selfie-Konfiguration) werden die Marker in allen untersuchten Bibliotheken sehr zuverlässig gefunden. Für seitliche Aufnahmen, wie sie in unserem Aufbau notwendig sind, zeigt sich eine große Herausforderung. Hier sehen wir die Notwendigkeit für das Anlernen der Modelle auch Daten aus den von uns verwendeten Blickrichtungen zu verwenden. Alternativ könnten die facial

landmarks aus Frontalaufnahmen lediglich für die Detektion der relevanten, aber nicht bewegten, Teile der Gesichtspartie verwendet werden, um die Textur zu extrahieren und anschließend texturbasierte Algorithmen für die Verarbeitung der seitlichen Aufnahmen zu verwenden.

8 Zusammenfassung und Ausblick

In dem Beitrag wurde ein „proof of concept“ zur Bewegungserkennung auf Basis von facial landmarks für den Einsatz in der Strahlentherapie vorgestellt. Basis dafür bildete die Beschreibung des Systemaufbaus und dessen Kalibrierung. Für die Bewegungserkennung wurden unterschiedliche Bibliotheken zur Detektion von facial landmarks verwendet. Dabei zeigen sich in dem verwendeten sehr einfachen Aufbau natürlich Verbesserungspotentiale in der Wahl der Hardware (Kameras) sowie in den verwendeten Marken (OpenCV, ArUco). Trotzdem konnten selbst mit diesem einfachen Aufbau die Größenordnung der benötigten Genauigkeit annähernd erreicht werden. Damit kann von einer prinzipiellen Machbarkeit des gewählten Lösungsansatzes ausgegangen werden.

9 Danksagung

Das Projekt „Increasing the precision of patient positioning in radiotherapy using positioning aids“ (imPRESSing) wird durch das Else Kröner-Fresenius-Zentrum für Digitale Gesundheit (TU-Dresden) gefördert.

Literatur

- [1] Debus, P., Bartenstein, J., Reiser, M. & Kuhn, F.-P. (2017): Duale Reihe, Radiologie. 4. vollständig überarbeitete Auflage. Thieme Stuttgart.
- [2] Schlegel, W., Karger, C. P. & Jäkel, O. (2018): Medizinische Physik Grundlagen – Bildgebung – Therapie – Technik. Springer Berlin Heidelberg.
- [3] VDI/VDE 2634 Blatt 1: Optical 3D measuring systems - Imaging systems with point-by-point probing (2002)
- [4] Dlib C++ Library: <http://dlib.net/>
- [5] opencv: OpenCV: Open Source Computer Vision Library <https://github.com/opencv/opencv>
- [6] mtcnn: <https://arxiv.org/abs/1604.02878> bzw. <https://doi.org/10.48550/arXiv.1604.02878>
- [7] mediapipe: <https://mediapipe.dev/>