

Bewegungsverfolgung für Präventions- und Rehabilitationsübungen: Evaluation des Kinect-Sensors und eines markerbasierten Systems

Andreas Wendleder*, Anne Grohnert*,
Stefan Klose**, Michael John**, Janka Siewert**

Technische Universität Berlin *

Fraunhofer Institut für Rechnerarchitektur und Softwaretechnik **

Zusammenfassung

Mit den Projekten Tivitrain und Reha@Home soll eine multimediale und motivierende Übungsplattform für gesundheitliche Prävention und Rehabilitation aufgebaut werden. Hierfür werden insbesondere Ansätze aus dem Serious Gaming mit einer medizinisch validen Bewegungsanalyse kombiniert. Als erster Schritt für die Auswahl der nötigen Sensorik wurde die Eignung und Präzision von marktüblichen Bewegungsverfolgungssystemen untersucht. Hierbei kamen jeweils ein markerbasiertes und ein markerloses System zum Einsatz. Zwei Übungen aus der Rehabilitation wurden simultan aufgenommen und miteinander verglichen. Basierend auf den Ergebnissen sollen in einem zweiten Schritt Verbesserungsvorschläge gegeben werden, wie solche Systeme an die qualitativ hohen Anforderungen von Ärzten und Therapeuten angepasst werden können.

1 Bewegungsverfolgung für gesundheitliche Prävention und Rehabilitation

Ein altersgerechtes Kraft- und Koordinationstraining wird im Alltag ständig benötigt, um einen Schutz für das gesamte Bewegungssystem zu bilden. Ein vorrangiges Prinzip des präventiven Gesundheitssports ist die Regelmäßigkeit seiner Realisierung sowie generell die Aufrechterhaltung eines aktiven Lebensstils. Kraft-, Koordinations- und Konditionstraining können helfen, die Bewegungs- und Funktionseinschränkung von älteren Arbeitnehmern und Menschen zu mindern (M. Kohli et al., 2000). Um die Patienten – nach der stationären Reha – langfristig an gesundheitssportliche Aktivitäten zu binden und sie zur Durchführung ergän-

zender Heimtrainingsprogramme zu bewegen, müssen spezielle Motivationsstrategien entwickelt werden (Pahmeier, 1999). Ein interaktives Übungsszenario, das im gesamten familiären Umfeld aktiv und spielerisch genutzt werden kann, könnte hier sehr erfolgreich sein.

Die Projekte TiviTrain¹ und Reha@Home² sollen dazu beitragen, diese Lücke zu schließen, indem die valide Erfassung und Ausführung von Bewegungsübungen im häuslichen Umfeld mittels innovativer Informations- und Kommunikationstechnik, Sensorik und Motivationsmethodik so unterstützt wird, dass ein nachhaltiger, medizinischen Kriterien genügender Übungserfolg gewährleistet wird. Für diesen Zweck ist es notwendig, bestehende Systeme und Sensoren hinsichtlich ihrer Qualität zu bewerten, um Wissen darüber zu erlangen, welche Komponenten und Eigenschaften verbessert werden müssen.

2 Anforderungen an die Bewegungsverfolgung aus therapeutischer Sicht

Rehabilitationsübungen sind in der Regel repetitive Bewegungen, die langsam und kontrolliert durchgeführt werden. Für die therapeutische Bewertung ist es notwendig, dass die Bewegungen begutachtet und bewertet werden. Während eines stationären oder ambulanten Klinikaufenthaltes erfolgt dies durch einen Therapeuten oder Arzt. Zur Erhöhung der Nachhaltigkeit ist es unerlässlich die Übungen auch nach dem stationären Aufenthalt im heimischen Umfeld mit einer gleichbleibend hohen Ausführungsqualität durchzuführen.

Im Folgenden werden die Anforderungen hinsichtlich der zu erzielenden Genauigkeit für die computergestützte Bewegungserfassung und -analyse beschrieben. Dafür werden zuerst die Bewegungen beschrieben, um davon ausgehend wichtige, zu berücksichtigende Bewegungsparameter abzuleiten. Die Bewegungsabläufe sind wie folgt charakterisiert:

Flankendehnung

Bei der Flankendehnung wird zuerst die Grundstellung eingenommen. Der Übende steht dabei im schulterbreiten Stand (1-2 schulterbreit), beide Arme liegen locker seitlich an den Oberschenkeln an. Danach wird ein Arm seitlich über den Kopf geführt wobei der Oberkörper in die Richtung gekippt wird, in die sich der Arm bewegt. Zu erkennende Fehlstellungen sind: Hohlkreuz, Rundrücken, Überstreckung der Halswirbelsäule, Rotation und Abknicken der Wirbelsäule, Fußstellung geringer als hüftbreit, Füße werden vom Boden abgehoben, falscher Arm wird über den Kopf geführt, Arm wird auf dem Kopf abgelegt.

¹ siehe <http://www.swt.tu-berlin.de/menue/forschung/projekte/tivitrain/>

² http://www.first.fraunhofer.de/en/news/press_releases/rehahome_das_rehabilitationsprogramm_fuer_zuhause/



Abb. 1 Bewegungsamplitude bei der Flankendehnung

Dynamische Brückenstellung

Die dynamische Brückenstellung beginnt, indem sich der Patient auf den Rücken legt, die Knie anwinkelt und die Handflächen und Fußflächen auf den Boden auflegt. Danach wird die Hüfte angehoben, bis Schultern, Hüfte und Knie eine Gerade bilden. Die Position wird circa fünf Sekunden gehalten. Danach wird die Hüfte wieder in die Ausgangsstellung abgesenkt. Mögliche Fehlstellungen sind: nicht symmetrische Grundstellung, Hohlkreuzbildung, keine angewinkelten Beine, Hüfte zu wenig oder zu stark gehoben, zu langes Verharren in der Brückenstellung.

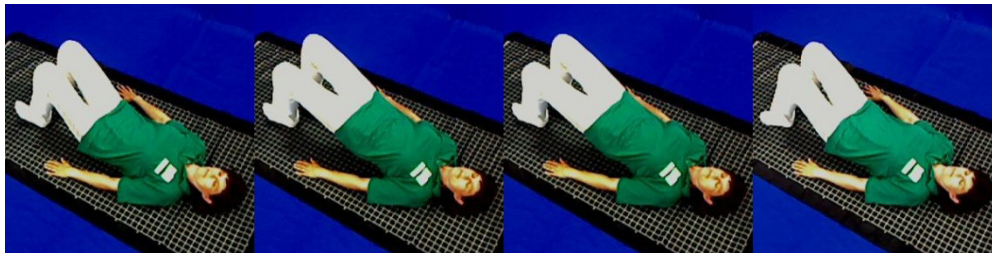


Abb. 2 Hüftanhebung bei dynamischer Brückenstellung

3 Evaluation der Bewegungsverfolgung

Die Gemeinsamkeit von markerlosen und markerbasierten Verfahren liegt in ihrer Aufgabe 3D-Bewegungsabläufe zu erfassen. Bei markerbasierten Verfahren wird ein IR-Kamerasystem aufgebaut, wobei je nach Komplexität der zu erfassenden Szene mindestens 6 Kameras um das Capturing Volume herum angeordnet sind. Am Probanden werden reflektierende Marker befestigt. Bekannte Systeme sind Imagemotion, das Vicon System sowie das hier verwendete Optitrack System.

Bei markerlosen Systemen erfolgt die Aufnahme mit einer oder mehreren RGB- oder Tiefenkameras. Um Bewegungen verfolgen zu können, muss im ersten Schritt der Mensch aus seiner Umgebung extrahiert werden. Aus dem Farbbild bzw. der Tiefenkarte kann dies über die Erkennung der Kontur und/oder Textur erfolgen (Viola & Jones, 2001; Wu & Nevatia, 2005; Gao et al., 2009; Zhu et al., 2010; Beleznai & Bischof, 2009). Anhand der 3D-

Punktwolke kann ein Skelett-Modell mit Hilfe geodätischer Abstände von Extrempunkten wie Hand, Fuß und Kopf gefunden (Plagemann et al., 2010; Ganapathi et al., 2010) oder anhand einer selbstorganisierenden Karte erstellt werden (Hacer et al., 2009). Für die hier beschriebene Evaluation der optischen Verfahren zur Bewegungsanalyse wurden jeweils ein markerloses und ein markerbasiertes Verfahren eingesetzt.

Das verwendete markerbasierte Motion-Capturing-System der Firma Optitrack besteht aus 6 Infrarotkameras vom Typ FLEX:V100R2, die in einem Kreis angeordnet sind. Der Proband trägt einen schwarzen enganliegenden Anzug, an dem 34 Marker angebracht sind. Die Kameras senden Infrarotstrahlen aus, die von den Markern reflektiert werden. Die Reflexionen sind für mehrere IR-Kameras sichtbar. Ist das System korrekt kalibriert, werden aus den aufgezeichneten Daten 3D-Positionen für alle Marker bestimmt. Aus diesen 3D-Punkten und unter Berücksichtigung eines biomechanischen Modells, kann in der Software Arena ein Skelett zusammengesetzt werden.

Als markerloses System kam die Kinect von PrimeSense und Microsoft zum Einsatz. Die Echtzeit-3D-Skeletterkennung der Kinect basiert auf Triangulation. Dazu sendet der Projektor ein für das menschliche Auge unsichtbares Infrarotpunktmuster aus. Die auf die Szene auftreffenden Punkte werden mit der Infrarotkamera erfasst. Das mittels Kinect erkannte Skelett besteht aus 15 Punkten (Kopf, Hals, Schultern, Ellenbogen, Hände, Torso (Bauchnabel), Hüfte, Knie und Füße). Zur Initialisierung des Skelettes muss der Nutzer die sogenannte Psi-Pose einnehmen. Nach erfolgreicher Initialisierung wird das erkannte Skelett in Echtzeit aktualisiert. Eine gleichzeitige Erkennung mehrerer Nutzer ist möglich.

Bei den Messungen sollte Robustheit und Genauigkeit der Bewegungsverfolgung für Rehabilitations-Übungen bestimmt werden. Dazu wurden mit beiden Systemen die Skelettanimationen beider Systeme mit einer Abtastrate von 30 Hz aufgezeichnet. Somit können die eingangs erwähnten Fehlstellungen und das Übungstempo (Geschwindigkeit der Extremitäten, Verharren an Umkehrpunkten) erkannt werden. Dafür wurden die Marker für das Optitrack System an der Versuchsperson angebracht und die Testperson vermessen.



Abb. 3 Vermessene Versuchsperson mit angebrachten Markern

Anschließend wurden die beschriebenen Übungen durchgeführt und simultan mit beiden Systemen aufgezeichnet. Die Skelettdaten des markerbasierten Verfahrens wurden im BVH-Format exportiert und zur weiteren Analyse mittels eines Skriptes in Blender3D importiert. Die Kinect-Daten wurden mit Hilfe einer selbstgeschriebenen Software aufgenommen, bearbeitet und einem Blender3D-kompatiblen Format gespeichert. In Blender wurden die beiden

Animationen visualisiert, so dass ein Physiotherapeut sofort erkennen kann, ob die Übungen richtig ausgeführt wurden.

Um die Datensätze der Systeme vergleichen zu können, wurde eine Software geschrieben, die beide Skelettanimationen normalisiert, Ausreißer entdeckt und korrigiert, sowie die Abweichungen beider Skelette zueinander berechnet. Anhand der Anzahl der Ausreißer und deren Abweichungen kann anschließend ein Rückschluss über die ursprünglich durchgeführte Bewegung getroffen werden. Somit kann auch bestimmt werden, welches der beiden Systeme die besseren Ergebnisse lieferte.

4 Ergebnisse der Vergleichsmessungen

Die angestellten Messungen zeigten, dass weniger die Genauigkeit, als vielmehr die Robustheit der Messung der kritische Punkt für die Erstellung von 3D-Skelettanimationen darstellt. Abgesehen von einzelnen Ausreißern aufgrund verdeckter Marker, stimmten beide Systeme gut überein, das heißt, dass die Abweichungen 5 cm nicht überschritten. Obwohl die Erfassung der Marker und ihre Umrechnung in 3D-Koordinaten zumeist hervorragend funktionierten, gab es bei der Zuordnung dieser Punkte zu korrespondierenden Skelett-Segmenten erhebliche Probleme. Sobald Marker für einige Frames nicht sichtbar waren, positionierte die Software die noch erkannten Marker fehlerhaft. Dies führte zu unerwarteten Verrenkungen in der Darstellung. Die markerlose Lösung der Kinect ist hier um einiges robuster, so dass das Skelett auch bei unzureichender Datenlage weiterhin plausibel dargestellt werden kann.

Bei der bodennahen Übung wurden die Grenzen beider Systeme erreicht, da entweder zu viele Marker verdeckt wurden bzw. die Positionen der Marker auf eine stehende Stellung optimiert sind.

Beim markerlosen System wirkt sich die geringe Auflösung des Tiefenbildes negativ aus. Ein im Raum stehender Mensch ist aufgrund des großen Abstandes zu den Wänden sehr gut separierbar. Durch den geringen Tiefenunterschied zwischen Mensch und Boden, verschmelzen Testperson und Boden miteinander, so dass das erzeugte Skelett fehlerbehaftet ist.

Vor- und Nachteile der beiden Systeme:

| | Kinect | Optitrack |
|-----------|--|--|
| Vorteile | Einfache Einrichtung, sehr gutes Preis-Leistungsverhältnis | Hohe Genauigkeit, schlechteres Preis-Leistungsverhältnis |
| Nachteile | Keine hohe Präzision der Messung | Hoher Einrichtungsaufwand aufgrund der Kalibrierung |

5 Zusammenfassung und Ausblick

In diesem Paper wurden ein markerloses und ein markerbasiertes Motion-Capturing-System verglichen. Es wurden eine im Stehen und eine in Bodennähe durchzuführende Übung betrachtet.

Zusammenfassend kann gesagt werden, dass die Skeletterkennung der Kinect einen großen Verbesserungsschritt in der markerlosen Erfassung von Bewegungen darstellt. Bis auf die erwähnten Probleme bei der Erfassung von bodennahen Übungen ist die Erkennungsrate hervorragend für die vorgestellten Rehabilitationsübungen geeignet.

Die markerbasierte Lösung bedingt einen höheren Einrichtungsaufwand, erfasst die Testperson allerdings auch sehr gut. Bei der Zuordnung von Markern zu korrespondierenden Skelettsegmenten traten allerdings teilweise Probleme auf. Hier wäre eine Optimierung der Software wünschenswert, so dass Ausreißer entdeckt und korrigiert werden.

In der Zukunft sollen auch weitere – insbesondere im Sitzen durchzuführende – Rehabilitationsübungen sowohl mit den beiden vorgestellten Systemen als auch mit weiterer körpernaher Inertialsensorik analysiert werden.

Literaturverzeichnis

- Beleznai, C. & Bischof, H. Fast Human Detection in Crowded Scenes by Contour Integration and Local Shape Estimation. In CVPR'09, 2009.
- Ganapathi, V., Plagemann, C., Koller, D. & Thrun, S. Real time motion capture using a single time-of-flight camera. In CVPR'10, 2010
- Gao, W., Ai, H. & Lao, S. Adaptive Contour Features in oriented granular space for human detection and Segmentation. In CVPR'09, 2009.
- Hacer, M., Böhme, M., Martinetz, T. & Barth, E. Self-Organizing Maps for Pose Estimation with a Time-of-Flight Camera. In Dynamic 3D Imaging – Workshop in Conjunction with DAGM, volume 5724 of Lecture Notes in Computer Science, pages 142 – 153, 2009.
- Kohli, M., Künemund, H., Motel, A. & Szydlik, M. Grunddaten zur Lebenssituation der 40-85-jährigen deutschen Bevölkerung. Ergebnisse des Alters-Survey. Berlin: Weißensee-Verlag, 2000.
- Pahmeier, I. Barrieren vor und Bindung an gesundheitssportliche Aktivität. Eine Rahmenkonzeption und empirische Untersuchung zu Merkmalen für Abbruch und Bindung im Gesundheitssport unter Berücksichtigung der sportbezogenen Selbstwirksamkeit. Habilitation. Universität Bayreuth, 1999.
- Plagemann, C., Ganapathi, V., Koller, D. & Thrun, S. Real-time Identification and Localization of Body Parts from Depth Images. In ICRA'10, 2010.
- Viola, P. & Jones, M. Rapid Object Detection Using a Boosted Cascade of Simple Features. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2001.
- Wu, B. & Nevatia, R. Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors. In CVPR'05, 2005.
- Zhu, Y., Dariush, B. & Fujimura, K. Kinematic self retargeting: A framework for human pose estimation. In Journal Computer Vision and Image Understanding, vol. 114, pp. 1362 – 1375, 2010.