

FiFi – Robuste Personenerkennung gestengesteuerter Fahrzeuge in der Intralogistik

FiFi – Robustness of Person Detection for Gesture Controlled Vehicles in Intralogistics

Andreas Trenkle
Jödis Rappl
Malte Grebe
Kai Furmans

Institut für Fördertechnik und Logistiksysteme (IFL)
Karlsruher Institut für Technologie (KIT)

Die Steuerung des Fahrerlosen Transportfahrzeuges „FiFi“ erfolgt berührungslos durch Gesten- und Personenerkennung basierend auf 3D-Daten der Umgebung. Die genutzten Verfahren zur Personenerkennung führen in einigen Fällen zur Falsch-Erkennung von Personen in Objekten. Das Paper beschreibt die Ursachen der Fehlerkennung und stellt die umgesetzten Lösungsansätze zur Vermeidung vor. Experimente bestätigen, dass die entwickelten Verfahren die Robustheit des Systems erhöhen.

[Gestensteuerung, Fehlerkennung, FTF, Personenverfolgung, Kinect]

The Automated Guided Vehicle (AGV) “FiFi” is controlled by recognition of gestures and persons based on 3D data of the environment. The algorithms of person recognition lead to false detection of persons in some cases. The paper describes error causes and presents some solutions. Experiments confirm that the developed method increases the robustness of the system.

[gesture control, false-detection, AGV, Person Following-Robot (PFR), Kinect]

1 MOTIVATION

Aufgrund der mangelnden Flexibilität von Automatisierungslösungen werden in der Intralogistik zahlreiche Transporte manuell bewerkstelligt. Die Bedienung manueller Transportmittel wie Handhubwagen, Elektro-Gabelhubwagen oder Kommissionierwagen erfordert physischen Kontakt zwischen Bediener und Gerät.

„FiFi“ (Abbildung 1) grenzt sich von bestehenden Transportsystemen durch seine Benutzerschnittstelle ab. Hier kommt ein 3D-Kamerasystem zum Einsatz, das die Umgebung und Bediener erfasst. Durch die Interpretation der Bewegungen des Bedieners leitet FiFi Befehle ab. FiFi entlastet den Bediener körperlich, indem das Fahrzeug ohne Berührung und menschlichen Kraftaufwand

folgt. Die Hände bleiben frei. Durch die eigenständige Entscheidung über Fahrbewegungen kann der Bediener sich vollständig auf die eigentliche Tätigkeit z.B. das Kommissionieren konzentrieren. Aufbau und Betriebsarten von FiFi werden in [Tre13] erläutert.



Abbildung 1. FiFi beim Transport eines Kleinladungsträgers

Die Erkennung von Personen und die korrekte Interpretation der Gesten sind Voraussetzung für die robuste Interaktion mit FiFi. Einerseits müssen Personen, welche mit FiFi interagieren, als solche erkannt werden. Andererseits darf FiFi Gegenstände in der Umgebung nicht als Personen erkennen. Die Interpretation der 3D-Daten ist anfällig für Fehlerkennungen, was Auswirkungen auf die Gestensteuerung hat und zu unerwünschtem Fehlverhalten von FiFi führen kann. Nachfolgend werden Lösungswege dargelegt, um die in FiFi implementierte Personenerkennung zu verbessern, im intralogistischen Umfeld eine hohe Erkennungsrate zu erzielen, sowie das Auftreten von Fehlerkennungen zu vermeiden.

Nach Erläuterung des Stands der Forschung zur Personenerkennung aus 3D-Daten werden in Abschnitt 3 die Ziele und die Vorgehensweise zur Regelfindung erklärt. In Abschnitt 4 werden die Regeln zur Vermeidung von Fehlerkennungen vorgestellt, bevor sie in Abschnitt 5 untersucht und bewertet werden. Abschließend erfolgen ei-

ne Zusammenfassung und ein Ausblick auf weiterführende Forschungsfragen.

2 GRUNDLAGEN

Die Skeletterkennung bezeichnet die Erstellung eines Mensch-Modells auf Basis von Sensordaten. Nachfolgend werden Arbeiten vorgestellt, die sich mit der Bewertung und der Verbesserung der Skeletterkennung befassen.

[Liv12] bewertet die Skeletterkennung der Kinect als eine vergleichsweise günstige Alternative zur Erkennung von Gesten. Die Bewegungen der Arme und Beine werden bei Einhaltung des idealen Abstandes zur Kinect von 1.2-3,5m korrekt lokalisiert. Probleme bestehen bei heller, direkter Lichteinstrahlung. Die in [Obd12] vorgestellten Ergebnisse zeigen den Vergleich der Skeletterkennung mit anderen, auch markerbasierten Bewegungserfassungssystemen. Betrachtet werden Bewegungsabläufe von Rehabilitationsübungen für ältere Menschen. Bei einfachen Körperbewegungen ist die Genauigkeit der Kinect vergleichbar mit anderen Bewegungserfassungssystemen. Bei der Verdeckung von Körperbereichen durch Überlappung oder Gegenstände verschlechtert sich die Genauigkeit. [Mar12] beschreibt ein System zur Erkennung von Ballett-Posen unter Nutzung der Skeletterkennung der Kinect und bestimmt die Güte des Systems anhand der Erkennungsrate der Posen.

Zur Erhöhung der Genauigkeit stellt [Yeu13] einen Ansatz vor, bei dem zwei Kinect-Kameras genutzt werden, um die Länge der Gliedmaßen eines Skelettes, sowie die Position einzelner Körperpunkte exakter zu bestimmen. Der Ansatz zeigt außerdem Vorteile bei der Erfassung von sich überlappenden Gliedmaßen.

Andere Ansätze verwenden neben einer Kinect weitere Kamerasysteme. So nutzt [Sus13] neben den RGB und Tiefendaten der Kinect eine Wärmebildkamera, während das in [Hos11] vorgestellte Verfahren einen Laserscanner mit einer Kinect zur Personenerkennung kombiniert.

Mittels Kinect-SDK für die Kinect [Msf14] werden sogenannte Skelettdaten des Nutzers berechnet. Diese sind über die Programmierschnittstelle *NUI Skeleton API* für eine weitere Verarbeitung nutzbar. Bei den Skelettdaten handelt es sich um 20 markante Körperpunkte, wie zum Beispiel Hüft-, Schulter-, Hand- oder Ellenbogen-Gelenke, die zu einem virtuellen Skelett des Nutzers verknüpft werden (vgl. Abbildung 2).

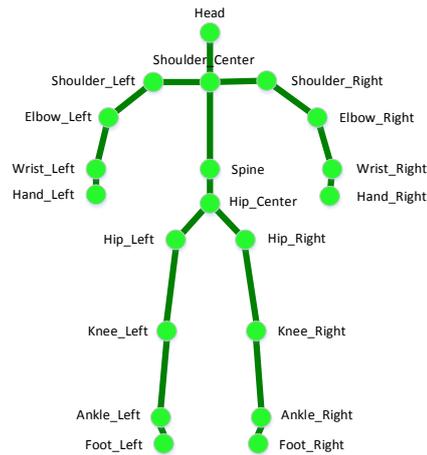


Abbildung 2. 20 Körperpunkte des Skeletts

Als *Skeletal Tracking* bezeichnet man die sequentielle Ermittlung und Nachverfolgung von Körperpunkten (en. *Joints*). Sobald der gesamte Körper von der Kamera erfasst wird, können alle 20 Körperpunkte erkannt werden. Je nach Qualität des erfassten Tiefenbildes werden die Körperpunkte eingeteilt in „lokalisiert“, „abgeleitet“ oder „nicht lokalisiert“. Kann ein Körperpunkt nicht direkt lokalisiert werden, ist es möglich, die Position aus der Lage anderer Körperpunkte „ abzuleiten“, wodurch die Genauigkeit stark abnehmen kann. Die ermittelten Skelettpunkte werden schließlich als 3D-Punkte im sogenannten *Skeleton Space* dargestellt. Die x-Achse erstreckt sich aus Sicht der Kamera nach links, die y-Achse nach oben und die z-Achse in die Blickrichtung der Kamera. In diesem kartesischen Raum bildet der Kinect-Sensor den Koordinatenursprung. Die Verbindungen zwischen den Körperpunkten werden als Knochen (en. *Bones*) bezeichnet.

Die in Abbildung 3 dargestellten Aufnahmen zeigen die von der Kamera erfassten Videodaten (rechts) und Tiefeninformationen (links). Die aus den Tiefendaten generierten Körperpunkte sind als hellgrüne Punkte dargestellt. Die Knochen werden als dunkelgrüne Verbindungslinien zwischen den Körperpunkten dargestellt.



Abbildung 3. Links: Tiefenbild mit eingezeichnetem Skelett, Rechts: Videodaten der Kamera

Kann die Position nur abgeleitet werden, so wird der Körperpunkt gelb markiert und der angrenzende Knochen als dünne helle Linie gekennzeichnet (vgl. Abbildung 4).



Abbildung 4. Abgeleitete Körperpunkte an Armen

3 ZIELSETZUNG UND VORGEHENSWEISE

Abbildung 5 zeigt die Vorgehensweise zur Vermeidung von Fehlerkennungen sowie die relevanten Vorarbeiten.

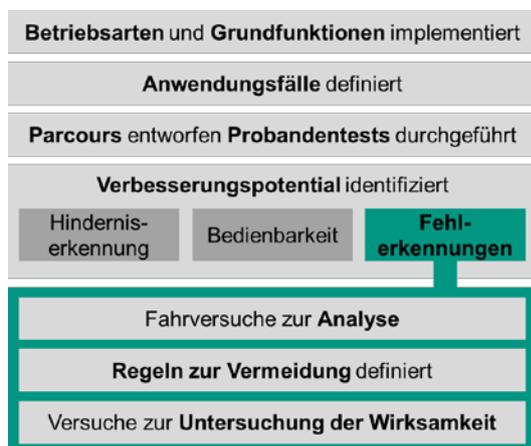


Abbildung 5. Vorgehensweise

Für FiFi wurden verschiedene Betriebsarten und Grundfunktionen implementiert, die in [Tre13] beschrieben werden. Aufbauend darauf wurden Anwendungsfälle definiert. Besonders sinnvoll ist der Einsatz in der manuellen Kommissionierung wo FiFi dem Bediener bei der Entnahme der Artikel durch die Regalgassen folgt. Nach Abarbeitung des Auftrages gibt der Bediener FiFi an eine optische Spur ab, das dann selbstständig zum nächstgelegenen Bereich fährt.

Um die Robustheit von FiFi zu bewerten, wurden Testparcours entworfen und Probandentests durchgeführt. Dabei zeigte sich Verbesserungspotential bei der Erkennung von Hindernissen, der Bedienbarkeit und der Vermeidung von Fehlerkennungen, auf die nachfolgend im Detail eingegangen wird.

Die Skeletterkennung der Kinect ist anfällig für Fehlerkennungen. Es wird zwischen zwei Formen der Fehlerkennung unterschieden:

Als *false positive* wird der Fall bezeichnet, wenn in einem Gegenstand im Sichtfeld der Kamera fälschlicherweise ein menschliches Skelett erkannt wird (vgl. Abbildung 6). Dies tritt meist dann auf, wenn die Form eines Objektes einer menschlichen Silhouette ähnelt oder eine

direkte Lichteinstrahlung auf die Kamera die Messungen der Tiefendaten behindert bzw. verfälscht. Das Pendant wird als *true positive* bezeichnet, d.h. eine Person befindet sich im Sichtfeld und ein korrektes Skelett wird dargestellt (vgl. Abbildung 3).

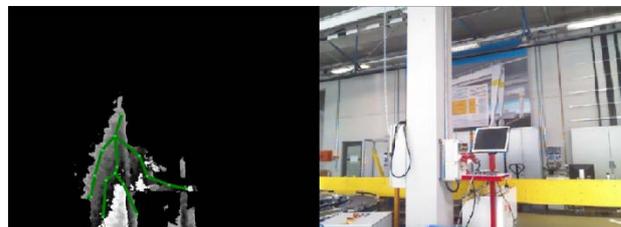


Abbildung 6. Beispiel einer false positive -Erkennung

Ein *false negative* tritt auf, sobald eine Person, die sich im Sichtfeld der Kinect-Kamera befindet, nicht als solche erkannt wird. Auch in diesem Fall kann die Lichteinstrahlung ein Grund für verfälschte Tiefendaten sein. Auch die Positionierung der Person zur Kamera hat Einfluss hierauf. Die Person kann einen zu großen oder zu geringen Abstand zur Kamera haben oder so zur Kinect ausgerichtet sein, dass aus den Tiefendaten keine Person erfasst werden kann. Das korrekte Pendant wird als *true negative* bezeichnet, d.h. es befindet sich keine Person im Sichtfeld und kein Gegenstand wird fälschlicherweise als Person erkannt.

Das Auftreten von *false positives* und *false negatives* reduziert die Robustheit des Systems und schränkt die Verfügbarkeit von FiFi ein.

Ziel der nachfolgend dargelegten Untersuchungen ist die Vermeidung von Fehlerkennung in Objekten (*false positives*). Eine Bedingung ist dabei, dass ein Anstieg von Fehlerkennungen realer Personen (*false negatives*) vermieden wird.

Zur Erreichung dieser Ziele wurden auf Grundlage von Fahrversuchen Regeln erarbeitet, die die Anordnung der Körperpunkte überprüfen. Sie zeigen an, ob es sich bei dem erkannten Skelett um eine Person handelt oder nicht. Die Wirksamkeit der implementierten Regeln wurde in Versuchen untersucht.

3.1 VERSUCHSFAHRTEN ZUR UNTERSUCHUNG DER FEHLERKENNUNGEN

Ausgangspunkt für die Aufstellung der Regeln sind *false positive* Erkennungen, welche im Verlauf von Versuchsfahrten in Logistik- und Produktionsbereichen aufgetreten sind.

Bei diesen Versuchsfahrten wurde beobachtet, dass Skelette in Gegenständen fälschlicherweise erkannt werden und im schlimmsten Fall auch zur Erkennung der Anmeldegeste führen. Ursache sind Messungenauigkeiten der Tiefendaten, wodurch eine Bewegung eines falschen Skelettes antizipiert und somit die Position der Körperpunkte verändert werden. Hierdurch „folgt“ FiFi

dem falsch erkannten Gegenstand und steht dem Bediener nicht mehr zur Verfügung. Deshalb sind die Fehlerkennung bereits vor einer möglichen Anmeldung aufzudecken.

Nachfolgend werden die bei den Versuchsfahrten gemachten Beobachtungen dargelegt, aus denen die im nächsten Abschnitt erläuterten Regeln abgeleitet wurden.

Körpergröße

Im Vergleich zu der Skelettdarstellung ausgewachsener Benutzer führen einige Fehlerkennungen zu Skeletten mit geringer oder sehr großer Körpergröße. Die Höhe der fälschlicherweise als Person erkannten Kiste in Abbildung 7 beträgt 60 cm.



Abbildung 7. Skelett mit geringer Körpergröße

Unnatürliche Körperhaltung

Die in Abbildung 8 getätigte Beobachtung einer Fehlerkennung zeigt ein für den Beobachter unnatürlich wirkendes Skelett, da die Körperpunkte in relativer Position zueinander nicht stimmig sind.

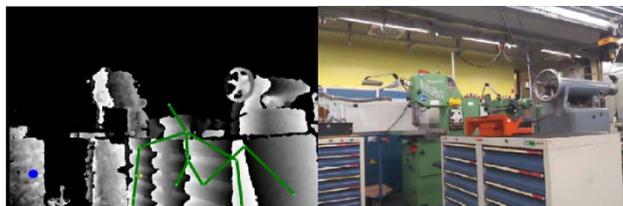


Abbildung 8. Ungewöhnliche Körperhaltung eines Skeletts

Liegeposition

Abbildung 9 zeigt ein Skelett, das eine liegende Haltung darstellt, die von Mitarbeitern im intralogistischen Umfeld üblicherweise nicht eingenommen werden sollte.



Abbildung 9. Körperhaltung ähnelt einer Liegeposition

Unnatürliche Arm- und Beinlänge

Das in Abbildung 10 erkannte Skelett weist eine unnatürliche Armlänge auf. Sowohl die Proportionen zum

Körper als auch die Differenz der Länge beider Arme sind auffällig.



Abbildung 10. Auffällige Differenz der Armlänge auf rechter und linker Seite des Skeletts

Viele abgeleitete Körperpunkte

Die in Abbildung 11 dargestellte Fehlerkennung zeigt ein Skelett bei dem 7 von 20 Körperpunkten nicht lokalisiert sondern abgeleitet wurden.



Abbildung 11. Skelett mit einer hohen Anzahl von abgeleiteten Körperpunkten (7/20)

Körpermittelpunkt zu hoch

Abbildung 12 zeigt ein Skelett, dessen Körpermittelpunkt sich auf einer unrealistischen Höhe befindet. Eine Bedienung von FiFi durch Personen, die auf einer Erhöhung stehen, kann ausgeschlossen werden.



Abbildung 12. Skelett mit unrealistischer Höhe der Hüfte

4 DEFINITION DER REGELN

Die Skeletterkennung des Microsoft-Frameworks ist auf privaten Einsatz ausgelegt, bei dem Bediener Körperhaltungen einnehmen, die im intralogistischen Umfeld ausgeschlossen werden können. Beispielsweise kann eine Interaktion mit FiFi im Liegen oder Sitzen ausgeschlossen werden. Weiterhin befinden sich im intralogistischen Umfeld erwachsene bzw. ausgewachsene Personen. Somit kann zwischen Fehlerkennungen im herkömmlichen Einsatzgebiet der Kinect-Kamera und Fehlerkennungen

im intralogistischen Einsatzbereich von FiFi unterschieden werden.

Der Verstoß einer Regel bewirkt, dass das erkannte Skelett als *false positive* eingestuft wird. Im Rückschluss darf im Idealfall ein *true positive* keine dieser Regeln verletzen. Für eine Regel wird vorausgesetzt, dass die genannten Körperpunkte lokalisiert werden und nicht abgeleitet sind – ausgenommen der Regel zur Anzahl der lokalisierten Körperpunkte. Dadurch wird verhindert, dass durch ein ungenaues Ableiten der Position eines oder mehrerer Körperpunkte Falschaussagen getroffen werden und es zu einem *false positive* oder *false negative* kommt.

Sobald ein für die Regel benötigter Körperpunkt nicht direkt lokalisiert werden kann, darf anhand dieser Regel keine Entscheidung getroffen werden, ob es sich um eine Fehlerkennung handelt oder nicht. In diesem Fall wird die Regel übersprungen und auf Grundlage der übrigen Regeln eine Entscheidung getroffen.

4.1 KÖRPERGRÖßE

Diese Regel beinhaltet die Überprüfung der Körpergröße eines erkannten Skeletts. Neben einer Mindestgröße wird eine Maximalgröße definiert. Gegen diese Regel wird verstoßen, sobald die gemessene Körpergröße außerhalb eines vorher festgelegten Intervalls liegt.

Gemessen wird die Körpergröße durch Addition des Abstandes der relevanten Körperpunkte im dreidimensionalen Raum. Hierfür werden acht Körperpunkte (siehe Abbildung 13) herangezogen (*Head, Shoulder_Center, Spine, Hip_Center, Hip_Left, Knee_Left, Ankle_Left* und *Foot_Left*). Alternativ können auch die Körperpunkte des rechten Beines genutzt werden, falls das linke Bein abgeleitete Körperpunkte beinhaltet. (*Head, Shoulder_Center, Spine, Hip_Center, Hip_Right, Knee_Right, Ankle_Right* und *Foot_Right*).

Da der Körperpunkt *Head* in der Mitte des Kopfes platziert ist, wird außerdem ein konstanter Wert von 10 Zentimetern addiert, um den Abstand der Kopfmitte bis zur Oberseite des Kopfes abzuschätzen. Die aktuell festgelegte Mindestgröße beträgt 120 Zentimeter, die Maximalgröße 225 Zentimeter. Dieses große Intervall ist gewählt, um der Nichterkennung realer Personen (*false negative*) durch Messungenauigkeiten entgegenzuwirken.

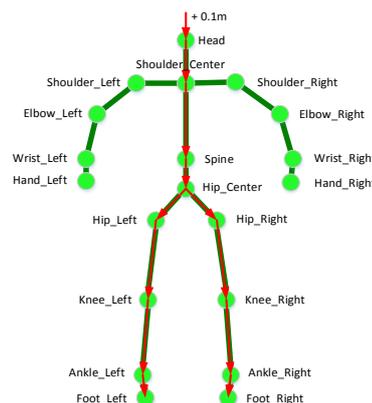


Abbildung 13. Körperpunkte zur Bestimmung der Körpergröße

4.2 ANZAHL LOKALISierter KÖRPERPUNKTE

Eine weitere Regel überprüft die Anzahl der Körperpunkte, welche von der Kinect direkt lokalisiert werden. Beobachtungen zeigen, dass die untere Körperhälfte des Benutzers nicht immer erfasst wird, da sich dieser z.B. zu nah an FiFi befinden kann. Aus diesem Grund werden für diese Regel ausschließlich die elf Körperpunkte der oberen Körperhälfte betrachtet (*Head, Shoulder_Center, Shoulder_Right, Shoulder_Left, Spine, Elbow_Left, Elbow_Right, Wrist_Right, Wrist_Left, Hand_Right, Hand_Left*).

Sobald die Anzahl der direkt lokalisierten Körperpunkte unterhalb von 9 liegt, wird dies als Verstoß gegen die Regel interpretiert und das erfasste Skelett als *false positive* interpretiert.

4.3 VERTIKALE ORDNUNG DER KÖRPERPUNKTE

Der Regel der vertikalen Ordnung liegt zugrunde, dass der Position eines Körperpunktes, bei aufrechter Körperhaltung, ein bestimmtes räumliches Verhältnis zu einem anderen Körperpunkt zugeordnet werden kann.

Da die Körperhaltung bei der Anmeldegeste bekannt ist, lassen sich daraus Regeln ableiten, wie dieses räumliche Verhältnis eines Körperpunktes zu einem anderen auszusehen hat. Betrachtet man die vertikale Achse, darf sich z.B. der rechte Fuß nur unterhalb der rechten Hüfte befinden. Befindet sich der rechte Fuß oberhalb der rechten Hüfte, so ist dies ein Indiz für ein *false positive*, da ein Verstoß gegen die Regel vorliegt.

4.4 HORIZONTALE ORDNUNG DER KÖRPERPUNKTE

Diese Regel ist von der Logik identisch mit der Vorherigen, allerdings wird nun die Horizontale betrachtet. Dementsprechend sollte sich aus der Sicht der Kinect beispielsweise die rechte Hand stets rechts von der linken Hand befinden. Befindet sich die rechte Hand links der linken Hand, so ist dies ein Indiz für ein *false positive*.

4.5 LIEGEPOSITION

Diese Regel soll alle Skelette ausschließen, deren Beinhaltung einer Liegeposition ähneln. Dabei wird ein Skelett als Fehlerkennung interpretiert, sobald dessen Beine eine bestimmte Körperhaltung einnehmen. Gegen diese Regel wird verstoßen, sobald entweder die Körperpunkte auf der rechten oder auf der linken Seite die Bedingungen erfüllen (vgl. Abbildung 14).

Der vertikale Abstand zwischen Hüfte und Knie (*Hip_Right* (*Hip_Left*), *Knee_Right* (*Knee_Left*)) und Knie und Sprunggelenk (*Knee_Right* (*Knee_Left*), *Ankle_Right* (*Ankle_Left*)) muss 25 cm unterschreiten. Außerdem muss für einen Verstoß der vertikale Abstand zwischen Hüfte und Sprunggelenk (*Hip_Right* (*Hip_Left*), *Ankle_Right* (*Ankle_Left*)) 50 cm unterschreiten.

Dies kann jedoch auch der Fall sein, wenn sich der Benutzer in einer Hocke befindet. Deshalb wird zusätzlich überprüft, ob der horizontale Abstand zwischen Hüfte und Sprunggelenk mindestens 60 cm überschreitet.

4.6 BEIN- UND ARMLÄNGE

Diese Regel kontrolliert das Längenverhältnis des rechten mit dem des linken Armes, sowie das des rechten mit dem des linken Beines. Dazu wird die Armlänge anhand des euklidischen Abstands entlang der Körperpunkte von der Schulter bis zur Hand (rechts: *Shoulder_Right*, *Elbow_Right*, *Wrist_Right*, *Hand_Right*) bzw. von Hüfte bis zum Fuß (rechts: *Hip_Right*, *Knee_Right*, *Ankle_Right*, *Foot_Right*) ermittelt.

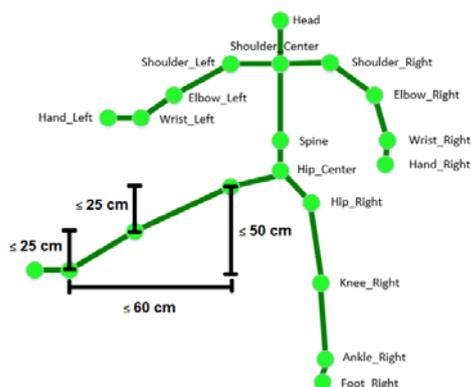


Abbildung 14. Beinhaltung zum Verstoß gegen die Regel „Liegeposition“

Überprüft wird, ob sich die Länge des rechten Arms nicht mehr als 15 cm von der des linken Arms unterscheidet. Analog werden die Längen der Beine überprüft. Unterscheiden sich die Bein- bzw. Armlängen um mehr als 15 cm, so wird das erkannte Skelett als Fehlerkennung interpretiert.

4.7 HÖHE DER HÜFTE

Diese Regel überprüft, ob sich der Benutzer vertikal in der Mitte des Sichtfeldes befindet. Da die Kamera den Koordinatenursprung bildet, wird festgelegt, dass sich der Benutzer bzw. der Körperpunkt der Hüfte (*Hip_Center*) nicht oberhalb eines Grenzwertes über dem Koordinatenursprung befinden darf. Dieser Grenzwert ist abhängig von der Anbringungshöhe der Kamera und muss dementsprechend bei Veränderungen angepasst werden.

5 EVALUIERUNG DER REGELN

Ziel des dargestellten Ansatzes ist die robustere Gestaltung der Personenerkennung von FiFi, ohne dabei die Interaktion zwischen FiFi und Benutzer zu beeinträchtigen. Um die Wirksamkeit zu zeigen, wurden die Regeln implementiert und durch weitere Versuchsfahrten ausgewertet.

Im Folgenden werden die Ergebnisse der Untersuchungen für die jeweiligen Regeln erläutert. Generell wurde während der Versuchsfahrten beobachtet, dass sowohl dunkle Kleidung als auch direkte Sonneneinstrahlung die Genauigkeit der Tiefendaten und damit die Skeletterkennung negativ beeinflussen.

5.1 EVALUIERUNG KÖRPERGRÖßE

Um eine Aussage über die Genauigkeit der Messwerte treffen zu können, wurden Personen mit unterschiedlicher Körpergröße gemessen. Steht die Person gerade und vollständig im Erfassungsbereich der Kamera, so ergibt die Größenmessung eine maximale Abweichung von ± 10 cm. Die Genauigkeit nimmt jedoch ab, sobald sich der Benutzer im Sichtfeld hin und her bewegt.

Ungenauigkeiten treten auf, wenn beispielsweise die Füße vom unteren Bildrand abgeschnitten werden. Dadurch werden die Körperpunkte der Sprunggelenke (*Ankle_Left* und *Ankle_Right*) fälschlicherweise nach oben versetzt. Bei dem in Abbildung 15 erfassten Benutzer lag daher die gemessene Körpergröße 20 cm unterhalb der tatsächlichen Körpergröße.

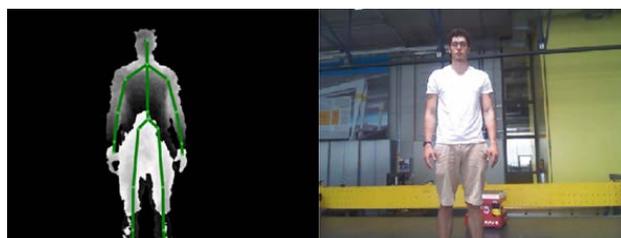


Abbildung 15. Körperpunkte des Unterkörpers zu hoch angesetzt

Der durch das Abschneiden der Füße maximal aufgetretene Messfehler beträgt 30 cm. Die Abweichungen

nach oben betragen maximal 10 cm. Auch wenn die Messung der Körpergröße demnach kein sehr exaktes Instrument ist, lassen sich eine Vielzahl der Fehlerkennungen verhindern. Geht man davon aus, dass die Körpergröße einer Person in der Intralogistik zwischen 150 cm und 215 cm liegt, so können alle Skelette ausgeschlossen werden, die kleiner als 120 cm und größer als 225 cm sind.

Für die Anwendung dieser Regel ist es erforderlich, dass der Körper vollständig lokalisiert werden kann. Deshalb sind Winkel und die Position, an der die Kamera an FiFi installiert ist relevant, um zu verhindern dass Körperpunkte (v.a. *Foot_Left* oder *Foot_Right*) abgeschnitten werden.

5.2 EVALUIERUNG ANZAHL LOKALISierter KÖRPERPUNKTE

Auch wenn der Bediener sich nur mit dem Oberkörper im Sichtbereich der Kamera befindet, lässt diese Regel eine Unterscheidung zwischen Person und Fehlerkennung zu. Beim Ausführen der Anmeldegeste fiel innerhalb der Versuchsfahrten die Anzahl der direkt lokalisierten Körperpunkte des Oberkörpers nicht unter 9 (von 11 möglichen). *False positives* wiesen jedoch überwiegend eine weitaus geringere Anzahl von lokalisierten Körperpunkten auf.

Falls durch die Position der Kamera versichert werden kann, dass sich die Person während der Anmeldung vollständig im Sichtfeld der Kinect befindet, ist zu überlegen, dass die Anzahl der direkt lokalisierten Körperpunkte des gesamten Körpers als Kriterium herangezogen werden.

Während den Versuchsfahrten wurde außerdem deutlich, dass bei vielen Fehlerkennungen der Körperpunkt des Kopfes (*Head*) nicht direkt lokalisiert werden konnte. Deshalb wird daraus eine weitere Regel abgeleitet, die die Lokalisierung des Kopfes prüft.

Darüber hinaus ist auffällig, dass gelegentlich die Körperpunkte der Hände, Handgelenke und Ellbogen lokalisiert werden konnten und gleichzeitig die Körperpunkte der Schultern nicht erkannt wurden. Bei der Anmeldung eines Benutzers tritt dieser Fall nicht ein. Dies spricht für einen weiteren Indikator einer Fehlerkennung.

5.3 EVALUIERUNG HORIZONTALE UND VERTIKALE ORDNUNG DER KÖRPERPUNKTE

Für die Anmeldung bei FiFi mittels Anmeldegeste gelten die Regeln der horizontalen und vertikalen Ordnung der Körperpunkte. Die Einführung einer vertikalen sowie horizontalen Ordnung der Körperpunkte verbessert die Robustheit des Systems.

Neben der horizontalen und vertikalen Position könnte in einem weiteren Schritt auch die Tiefenposition (z-Koordinate) überprüft werden. Häufig wird ein Skelett

in Säulen oder gestapelten Boxen erkannt. Im Gegensatz zu menschlichen Beinen zeigen die Beine dieser *false positives* keinen leeren Raum zwischen den Knien auf. Durch die Vorschrift, dass zwischen den Knien die z-Koordinate abrupt ansteigen muss, könnten Fehlerkennungen in solchen Gegenständen eventuell vermieden werden.

5.4 EVALUIERUNG LIEGEPOSITION

Die Regel zur Überprüfung, ob sich ein Skelett in einer Liegeposition befindet, verhilft im Vergleich zu den anderen Regeln eher selten zur Aufdeckung eines *false positives*. Dies lässt sich jedoch damit erklären, dass diese Regel einen sehr speziellen Fall betrachtet.

5.5 EVALUIERUNG HÖHE DER HÜFTE

Bei dieser Regel ist zu beachten, dass sich die formulierten Grenzwerte für die Abweichungen mit dem Standort der Kinect verändern, da die Kinect den Koordinatenursprung bildet. Die während den Versuchsfahrten gesetzten Grenzwerte waren auf die Position der Kinect am Versuchsfahrzeug angepasst. Die Ergebnisse zeigen, dass die Anwendung dieser Regel zur Aufdeckung von *false positives* verhilft. So konnten mehrere Skelette ausgeschlossen werden, deren Körperpunkte der Hüfte (*Hip_Center*) sich zu hoch im Sichtfeld befanden.

5.6 EVALUIERUNG BEIN- UND ARMLÄNGE

Die Überprüfung der Bein- und Armlänge erweist sich ebenfalls als sinnvoll. Weiterhin ist vorstellbar, die Überprüfung der Länge einzelner Knochen oder ganzer Körperabschnitte auszuweiten.

6 ZUSAMMENFASSUNG UND AUSBLICK

Durch die Umsetzung der Regeln konnten bei Versuchsfahrten 29 von 30 (96,7%) Fehlerkennungen als *false positive* aufgedeckt werden. Weiterhin wurde überprüft, ob die Regeln nicht auch potentielle Benutzer als Fehlerkennung einordnen. Dazu wurden in einem ersten Versuch Probanden mit unterschiedlicher Körpergröße und Körperform aufgefordert sich nacheinander in das Sichtfeld der Kinect zu stellen, sich per Anmeldegeste anzumelden und das Sichtfeld wieder zu verlassen. Dabei konnten 38 von 40 Anmeldungen erfolgreich durchgeführt werden. Dies entspricht einer Erfolgsquote von 95%. Die Ursachen für die nicht-Erkennung von zwei Anmeldeversuchen waren dabei verfälschte Tiefendaten, wodurch die Skelette nicht korrekt dargestellt und damit gegen mehrere Regeln verstoßen wurde oder eine inkorrekt ausgeführte Anmeldegeste des Benutzers.

Die Ergebnisse zeigen, dass die Anwendung einfacher Regeln einen positiven Einfluss auf die Vermeidung von Fehlerkennungen hat. Zum einen können Fehlerken-

nungen von Personen und Objekten vermieden werden und zum anderen werden Personen zuverlässiger erkannt.

Um eine hohe Erkennungsrate von Personen zu gewährleisten, signalisiert FiFi Fehlerkennungen mittels LED-Signalen. Damit hat der Bediener die Möglichkeit durch die Änderung seiner Position die Wahrscheinlichkeit einer Erkennung zu erhöhen. So kann der Benutzer im Falle eines *false negatives* darauf aufmerksam gemacht werden, dass eventuell aufgrund äußerer Bedingungen (z.B. schwarze Kleidung und direkte Lichteinstrahlung) die Tiefendatenerfassung beeinträchtigt wird.

Neben Vermeidung von Fehlerkennungen ist auch das Wiedererkennen von Personen von Nutzen, um die Robustheit zu erhöhen und die Bedienbarkeit von FiFi zu verbessern. In dem in Abbildung 15 dargestellten Szenario biegt der Bediener hinter einem Objekt ab und verlässt damit den Sichtbereich von FiFi. Aktuell meldet FiFi den Sichtverlust mit einem akustischen Signal und der Bediener muss sich erneut anmelden.

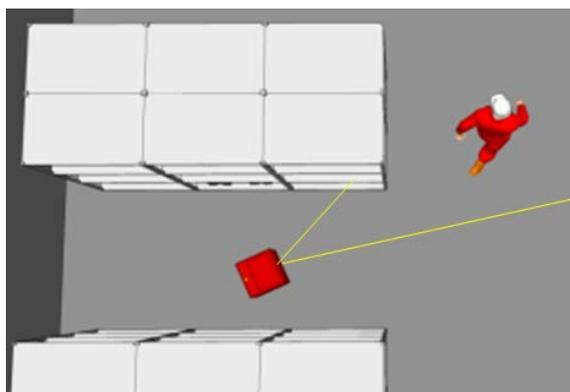


Abbildung 15: Person läuft aus Sichtbereich von FiFi heraus

Ziel der aktuellen Forschungsarbeiten ist die Nachverfolgung des Laufweges und die Wiedererkennung der Person. Hierzu müssen Laufweg des Bediener erfasst und bei Sichtverlust verfolgt werden. Da sich nach dem Abbiegen mehrere Personen im Sichtbereich befinden können, muss FiFi den zuvor angemeldeten Bediener anhand zuvor erfasster Merkmale wiedererkennen. Hierdurch ist eine erneute Anmeldung nicht erforderlich wodurch Zeit gespart und die Bedienbarkeit verbessert werden soll.

LITERATUR

- [Tre13] Trenkle, Andreas; Seibold, Zázilia; Stoll, Thomas; Furmans, Kai: *FiFi – Steuerung eines FTF durch Gesten- und Personenerkennung*. In Logistics Journal: Proceedings. 2013.
- [Liv12] Livingston, Mark A.; Sebastian, Jay; Zhuming, Ai; Decker, Jonathan W.: *Performance Measurements for the Microsoft Kinect Skeleton*. In Virtual Reality Short Papers and Posters (VRW), 2012 IEEE , vol., no., pp.119,120, 4-8 March 2012
- [Obd12] Obdrzalek, S.; Kurillo, G.; Ofli, F.; Bajcsy, R.; Seto, E.; Jimison, H.; Pavel, M.: *Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population*. In Proc. IEEE EMBC. 2012. S. 1188–1193.
- [Mar12] Marquardt, Zoe; Beira, João; Em, Natalia; Paiva, Isabel; Kox, Sebastian: *Super Mirror: a kinect interface for ballet dancers*. In CHI '12 Extended Abstracts on Human Factors in Computing Systems. 2012.
- [Yeu13] Yeung, Kwok-Yun; Kwok, Tsz-Ho; Wang, Charlie C. L.: *Improved Skeleton Tracking by Duplex Kinects: A Practical Approach for Real-Time Applications*. In J. Comput. Inf. Sci. Eng. 2013.
- [Sus13] Susperregi, L.; Sierra, B.; Castrillón, M.; Lorenzo, J.; Martínez-Otzeta, J. M., & Lazkano, E.: *On the Use of a Low-Cost Thermal Sensor to Improve Kinect People Detection in a Mobile Robot*. In Sensors 13, no. 11. S. 14687-14713.
- [Hos11] Hoshino, Fumiaki; Morioka, Kazuyuki: *Human Following Robot based on Control of Particle Distribution with Integrated Range Sensors*. In System Integration (SII). S. 212-217. 2011.
- [Msf14] Onlinequelle: *Kinect For Windows SDK Documentation*: <http://msdn.microsoft.com/en-us/library/hh855355.aspx>, abgerufen am 12. Mai 2014