

# Raumbasierte Erkennung von Stürzen mittels Tiefensensor und integrierter Verarbeitung

## Room based Fall Detection System using Depth Sensor and integrated Analysis

M. Schwaab<sup>1</sup>, L. Ketterer<sup>2</sup>, M. Wagner<sup>2</sup>, S. Maier<sup>2</sup>

<sup>1</sup> Hahn-Schickard-Gesellschaft für angewandte Forschung e.V., Wilhelm-Schickard-Str. 10, 78052 Villingen-Schwenningen, Manuel.Schwaab@Hahn-Schickard.de

<sup>2</sup> KUNDO xT GmbH, 78112 St. Georgen

### Kurzfassung

Ein raumbasiertes Sensorsystem wurde entwickelt mit dem der Sturz von (älteren) Menschen erkannt werden kann. Als Sensorik wird dabei eine Tiefen- bzw. 3D-Kamera verwendet welche im zu überwachenden Raum angebracht ist. Die Auswertung der Sensordaten findet dazu in Echtzeit in einem in dem Sensor integrierten Einplatinencomputer (Raspberry-PI 2) statt. Zur Verarbeitung gehört die automatische Kompensation der Neigung über einen Beschleunigungssensor, das Tracking der Person(en) sowie die Detektion von Stürzen. Durch die Auswertung im Sensor müssen keine Sensordaten übertragen werden, was natürlich auch aus datenschutztechnischer Sicht interessant ist. Die Algorithmen wurden anhand zahlreicher Testdaten optimiert bzw. evaluiert. Die ersten Ergebnisse zeigen, dass mit dem entwickelten System eine zuverlässige Sturzerkennung möglich ist.

### Abstract

A room based sensor system was developed to detect the fall of (elderly) people. For the monitoring a depth or 3D camera placed in the room is used. The analysis of the sensor measurement is carried out in real time on a single-board computer (Raspberry-PI 2) which is integrated into the sensor. The processing steps include the compensation of the sensor inclination using an accelerometer, the tracking of the person within the room as well as the detection of falls. The advantage of the integrated processing is that no sensor measurement have to be transmitted which can be very important in terms of data privacy. The algorithms were optimized and tested with several different recorded data sets. First results indicate that the system can detect falls very reliably.

## 1 Einleitung

In Zeiten steigender Lebenserwartungen und Senioren, die lange körperlich und geistig fit sind, wollen viele Menschen möglichst lange zuhause, in der vertrauten Umgebung wohnen, unabhängig vom gesundheitlichen Zustand. Die größte Angst, die die Lebenswirklichkeit der Senioren prägt, ist ein Ereignis, das dieses Wohnen zuhause nicht mehr ermöglicht. Zu diesen Ereignissen zählen Stürze, die nicht selten in eine Pflegebedürftigkeit münden. Dabei gibt es technische Hilfsmittel, die eine rasche Hilfe ermöglichen und damit den gravierenden Folgen vorbeugen (siehe auch [1] für eine ausführlichere Diskussion).

Mit einer automatischen Sturzerkennung kann älteren Menschen bei einem Sturz schnell Hilfe geschickt werden. Dies kann die Sicherheit steigern und ermöglicht ggf. auch ein längeres Leben in der eigenen Wohnung.

Im Gegensatz zu Sensoren, die im Boden eingebaut werden, hat das hier vorgestellte Raumbasierte System den Vorteil, dass keine größeren baulichen Veränderungen vorgenommen werden müssen. Werden körperbasierte Systeme betrachtet, liegt der Vorzug dieses Systems da-

rin, dass es nicht notwendig ist Sensoren am Körper zu tragen.

## 2 Verwandte Arbeiten

Die zunehmende Wichtigkeit dieser Problemstellung sowie die Verfügbarkeit von neuen bzw. besseren Sensortechnologien hat zur Entwicklung einer Vielzahl von verschiedenen Ansätzen und Verfahren zur automatischen Sturzerkennung geführt. Unterschieden werden dabei vor allem am Körper getragene Systeme, hauptsächlich mittels Inertialsensoren, von den raumbasierten Systemen z.B. mit Kameras, Mikrofonen oder Drucksensoren im Fußboden. Eine umfangreiche Übersicht zu den veröffentlichten Ansätzen ist in [2] zu finden. Einige Arbeiten mit ähnlichen Ansätzen, d.h. raumbasierte bzw. berührungslose Sturzerkennung mit Tiefenkameras werden im Folgenden noch etwas genauer beschrieben.

Eine Sturzerkennung mit einer an der Decke angebrachten, vertikal nach unten ausgerichteten Kinect Kamera wurde in [3] vorgestellt. Durch diese Anbringung und den relativ geringen Öffnungswinkel der Kinect ist der Erfassungsbereich, je nach Deckenhöhe, aber relativ eingeschränkt. Ein Ansatz über maschinelles Lernen wurde in [4] präsentiert. Dazu werden Merkmale die Formen der

menschlichen Silhouette entsprechen extrahiert. Anhand von aufgezeichneten Testdaten mit verschiedenen Aktionen wurde ein neuronales Netz trainiert aufgrund dann die Sturzklassifizierung erfolgt. Ebenfalls mit den Tiefendaten der Kinect wurde in [5] ein Personentracking implementiert. Für die Klassifikation wird ein Entscheidungsbaum verwendet, welcher anhand von verschiedenen Merkmalen, unter anderem der Bodenbereich der von der Person eingenommen wird oder auch der vertikalen Geschwindigkeit und Beschleunigung, entscheidet, ob ein Sturz vorlag oder nicht.

Bezgl. der verschiedenen, verwendeten Verarbeitungsschritte konnte dabei auf klassische Verfahren aus dem Bereich des maschinellen Sehens (siehe z.B. [6]) zurückgegriffen werden.

### 3 Sensorsystem

Die Sturzerkennung basiert auf einem Tiefensensor. Der verwendete Sensor bestimmt die Tiefe in dem die Laufzeit von Infrarotlicht gemessen wird (dadurch funktioniert er auch bei Dunkelheit). Der Sensor ist kalibriert und liefert entsprechend des Kameramodells eine 3D Koordinate (relativ zum Sensor) für jeden Bildpunkt.

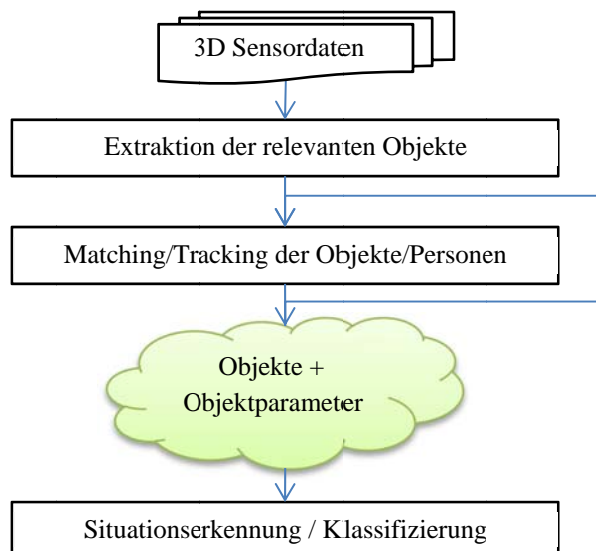
Die Sensorkamera hat einen relativ großen Öffnungswinkel von 90°, so dass der nicht erfasste Bereich nahe des Sensors relativ klein ist. Als Updaterate wurden 15 Hz verwendet, was für die Geschwindigkeit der menschlichen Bewegung absolut ausreichend war. Daneben bietet der Sensor noch andere Einstellungsmöglichkeiten, die allerdings für das Ergebnis nur eine untergeordnete Relevanz haben. Die Tiefensensordaten können über einen LAN Anschluss ausgelesen werden.

An diesen Sensor wurde ein Raspberry PI mit Linux Betriebssystem angebaut. Dieser liest die Daten des Sensors aus, speichert diese bzw. streamt sie weiter an einen PC. Außerdem wurde ein Neigungssensor (3-Achs Beschleunigungssensor über SPI, mit welchem die Neigung über die Gravitation bestimmt werden kann) in den 3D-Sensor integriert, um automatisch die Neigung des Sensors bestimmen zu können. Dies ist wichtig, dass der Sensor je nach Wohnung und Abdeckungsbereich auf unterschiedlicher Höhe und Neigung angebracht werden kann. Für eine möglichst gute Abdeckung bzw. Sichtbereich ist im Allgemeinen eine relativ hohe Anbringung unter der Decke sinnvoll.

Der Erfassungsbereich des Sensors in der Tiefe beträgt bis zu 10m – durch die Auflösung von 160 x 120 Bildpunkten sowie die größere Ungenauigkeit und Störanfälligkeit bei größeren Entfernungen ist eine zuverlässige Sturzerkennung aber nur in einem Bereich bis etwa 5m möglich.

### 4 Auswertung

Um die gesuchten Informationen aus den Sensormessungen extrahieren zu können, müssen diese verarbeitet werden. Die groben Schritte des Auswertalgorithmus welche für jedes aufgenommene Tiefenbild durchgeführt werden, sind in (Bild 1) skizziert.



**Bild 1** Vereinfachte Ablauf Skizze mit den groben Schritten des Algorithmus zur Auswertung der 3D-Sensordaten.

Wie zu sehen werden aus den vom Tiefensensor kommenden 3D Daten relevante Objekte extrahiert, welche dann über die Zeit getrackt werden und für welche eine Zustandsklassifikation durchgeführt wird. Für das Tracking werden die Objekte inklusive aussagekräftigen Parameter (z.B. Größe, Geschwindigkeit, ...) in einer Liste gespeichert. Die einzelnen Schritte werden im Folgenden genauer beschrieben.

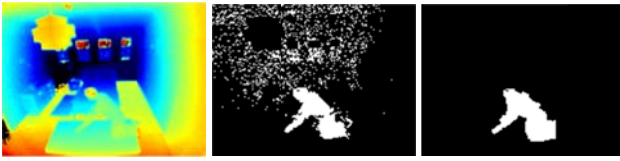
#### 4.1 Erkennung von relevanten Objekten

Ein generelles Szenenverständnis ist extrem schwierig und aufwendig. Deshalb werden nur Objekte berücksichtigt, die sich bewegen (im Normalfall Personen) bzw. solche, die ihren Platz verändern.

Dementsprechend wird zunächst der statische Hintergrund der vom Sensor wahrgenommenen Szene entfernt. Dazu wird anhand der ersten aufgezeichneten Bildern ein Hintergrund-Tiefenbild erstellt, wobei das Rauschen entsprechend berücksichtigt bzw. herausgefiltert wird. Dabei kann es auch sein, dass für einen Bildpunkt keine Hintergrundtiefe bestimmt werden kann. Wie im nächsten Abschnitt skizziert, kann sich dieses Hintergrund-Tiefenbild aber durchaus mit der Zeit ändern, z.B. wenn Möbel verstellt werden. Im Vergleich zum Maschinellen Sehen mit einer Farbkamera ist das Problem der Hintergrund-Subtraktion dabei etwas vereinfacht, da sich durch die aktive Beleuchtung die Helligkeit nicht mit der Zeit ändert. Das Hintergrund-Tiefenbild wird nun von den aufgenommenen Frames subtrahiert, so dass man eine Differenziefenkarte erhält. Bezgl. der Entscheidung, ob ein Pixel zum Hinter- oder Vordergrund gehört, sind verschiedene Ansätze bekannt. Ein einfacher globaler Grenzwert basierter Ansatz hat sich hier aber als ausreichend robust und am wenigsten rechenintensiv erwiesen. Der Grenzwert wurde dabei entsprechend dem Rauschen verschiedener Tiefenmessungen bestimmt.

Um die Fehldetektionen/Rauschen sowie Löcher in Objekten aus dem Vordergrundbild zu entfernen, wird danach noch eine gewöhnliche morphologische Erosion

durchgeführt, gefolgt von einer Dilatation, welche allerdings die Tiefeninformationen berücksichtigt (siehe **Bild 2**).



**Bild 2** Links: farblich codiertes Tiefenbild aufgenommen in einer Testwohnung. Mitte: Bewegungskarte nach der Hintergrund-Subtraktion. Rechts: Ergebnis nach der Erosion und der Dilatation (nur die Person ist noch übrig).

Danach werden Vordergrundpixel in zusammengehörenden Clustern unterteilt. Um zu erkennen, ob zwei benachbarte Pixel zum selben Cluster gehören, wird der Abstand der zugehörigen 3D-Punkte verglichen. Liegt dieser unter einem vorgegebenen Schwellwert, werden diese zum gleichen Objekt gehörig betrachtet. Dementsprechend wird ein sogenannter rekursiver 4-connected Floodfill Algorithmus eingesetzt, um zusammenhängende Cluster zu extrahieren.

## 4.2 Tracking von Personen/Objekten

Anhand von statistischen Merkmalen der Cluster bzw. der dazugehörigen Messungen können diese über die Zeit getrackt werden. Dazu werden die extrahierten Objekte aus den momentanen Messdaten mit zuvor erkannten Objekten verglichen und für jeden Vergleich eine Distanzmetrik bzw. einen Score ermittelt. Die Distanzmetrik besteht dabei hauptsächlich in der 3D-Position und den in den Tiefendaten sichtbaren realen Abmessungen des Objekts (unter der Annahme, dass 15fps die Änderungen durch die Bewegung relativ gering sind).

Das Objekt mit dem jeweils besten Score, welcher allerdings über einem bestimmten Grenzwert liegen muss, wird dann als die korrekte Assoziation (d.h. gleiches Objekt) verwendet.

Um bewegte Objekte besser tracken zu können, wird außerdem die Geschwindigkeit ermittelt, mit der sich das Objekt bewegt. Für den beschriebenen Vergleich kann somit die Position der zuvor erkannten Objekte näherungsweise vorhergesagt werden. Befindet sich ein erkanntes Cluster an der vorhergesagten Stelle, so ist es sehr wahrscheinlich, dass dieser zu dem entsprechenden Objekt gehört.

Durch das beschriebene Tracking kann die Bewegung von Personen über die Zeit verfolgt werden. Aber auch wie lange z.B. ein umgestellter Stuhl sich schon an der neuen Position befindet.

## 4.3 Klassifizierung

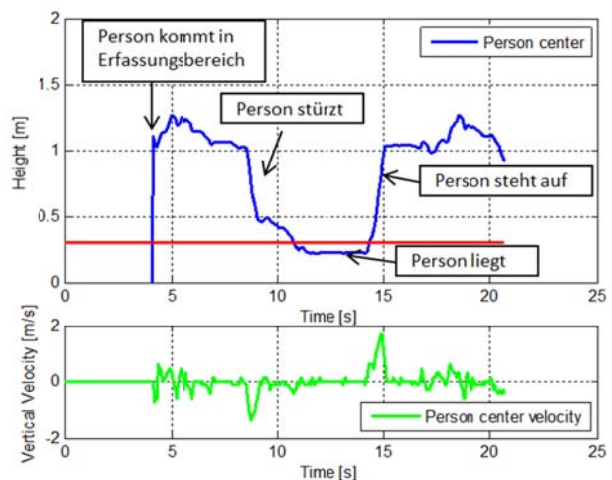
Die Klassifizierung der gefundenen und getrackten Objekte erfolgt anhand verschiedener Parameter. Bspw. muss sich ein erkanntes Objekt bewegt haben und muss eine bestimmte Mindestgröße aufweisen um als Person

klassifiziert zu werden (um z.B. die Verwechslung mit Haustieren zu verhindern).

Des Weiteren wird eine Klassifizierung des Zustands der Person durchgeführt. Folgende Zustände werden dabei ermittelt:

1. Stehend/gehend
2. Sitzend
3. Liegend (erhöht), z.B. Bett, Sofa, ...
4. Person am Boden
5. Sturz

Das Szenario, das beim Putzen oder Hervorholen eines Objekts unter dem Bett, bei welchem eine Person ebenfalls sehr nah am Boden, ist kann anhand einer Messung nicht zuverlässig von einer gestürzten Person am Boden unterschieden werden. Deshalb gibt es denn Zustand 4 bei welchem ein Alarm nur ausgelöst wird wenn die Person nach einer bestimmten Zeit nicht wieder aufgestanden ist. Für die direkte Sturzerkennung (Zustand Nummer 5) wird noch zusätzlich die Sturzgeschwindigkeit wie in (**Bild 3**) dargestellt berücksichtigt.



**Bild 3** Höhenverlauf des getrackten Körpermittelpunkts für die in Abbildung 1 dargestellte Szene. Ein wichtiges Kennzeichen für einen Sturz ist, dass der Körpermittelpunkt unter einem bestimmten Wert liegt (rote Linie). Aber auch die Geschwindigkeit, mit der man auf den Boden geht sowie die Dauer des Liegens kann als Merkmal genutzt werden.

Kann eine Person nicht mehr erfolgreich getrackt werden, wird überprüft, ob beim letzten Detektionszeitpunkt die Person am Rande des Erfassungsbereichs des Sensors war. Ist dies der Fall kann davon ausgegangen werden, dass die Person den Raum verlassen hat. Dies kann bei Bedarf ebenfalls an die Zentrale gemeldet werden.

Unbewegte Objekte werden im Allgemeinen nach einem gewissen Zeitraum verworfen und in den Hintergrund verschoben, so dass z.B. ein umgestellter Stuhl nach einer gewissen Zeit nicht mehr als Objekt getrackt wird.

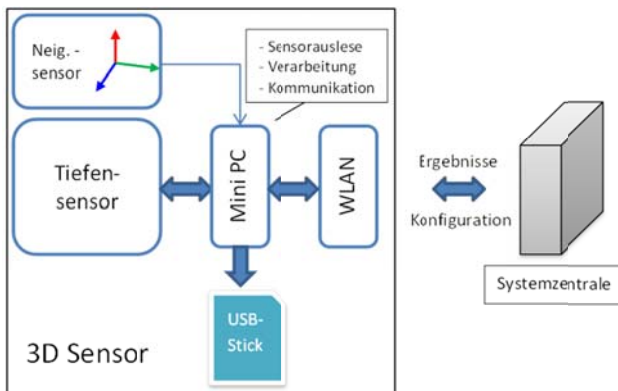
## 5 Demonstrator

Ein Demonstrator wurde entwickelt, bei dem die Verarbeitung in Realzeit komplett auf dem Sensor bzw. in dem

Sensor integrierten Raspberry PI 2 abläuft. Das bedeutet, dass die Daten im Sensor ausgewertet und lediglich die Ergebnisse an die Zentrale übertragen werden. Die aufgenommenen Tiefenbilder werden dementsprechend sofort nach der Verarbeitung wieder verworfen. Dadurch reduziert sich die Datenmenge, die übertragen werden muss – was natürlich auch in Bezug auf den Datenschutz vorteilhaft ist.

## 5.1 Hardware

Die Demonstrator Hardware setzt auf das für die Messungen verwendete Setup in Abschnitt 3 auf. Neben dem Tiefensensor besteht das Sensorsystem vor allem aus dem Raspberry PI. Durch den Beschleunigungssensor wird anhand der Gravitation die Neigung berechnet und entsprechend kompensiert. Über den am Raspberry PI angeschlossenen WLAN-Stick wird die Kommunikation mit der Systemzentrale ermöglicht. Auf einem USB-Stick werden darüber hinaus Logdaten sowie Messdaten kritischer Situationen gespeichert (für Entwicklungszwecke). Das beschriebene Konzept ist in **(Bild 4)** skizziert.



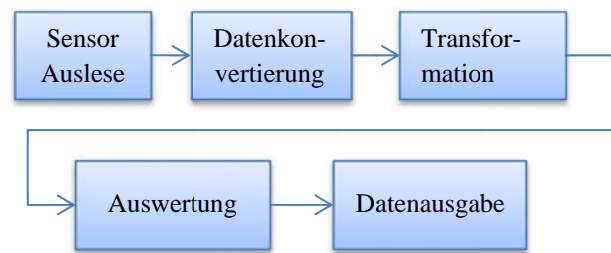
**Bild 4:** Hardware Konzept des Demonstrators für die Sturzerkennung.

Abgesehen von der notwendigen Stromversorgung kann das Demonstrator System, unter anderem durch die automatische Neigungskompensation, entsprechend der optimalen Abdeckung relativ beliebig in der Wohnung angebracht werden.

## 5.2 Software

Die Software welche auf dem Raspberry PI implementiert wurde beinhaltet vor allem die Portierung der auf dem PC entwickelten Algorithmen sowie die Integration des Sensors in das Gesamtsystem.

Für die Implementierung der Verarbeitungsschritte wurde, wie in **(Bild 5)** gezeigt, eine Pipeline Architektur verwendet.



**Bild 5** Verarbeitungspipeline, die auf dem in den Sensor integrierten Mini-PC implementiert ist.

Die Datenkonvertierung wandelt das proprietäre Format der 3D-Sensor-Daten in ein schnittstellenkonformes Format um. Das heißt nur die ersten Stufen sind 3D-sensor spezifisch. Da der Sensor (relativ) beliebig in der Wohnung angebracht werden kann, muss die Höhe bzw. Neigung kompensiert werden (der Algorithmus geht davon aus, dass die Koordinaten von Fußbodenmesswerten eine Höhe von 0 haben). Deshalb müssen die 3D-Punkte entsprechend rotiert und verschoben werden. Die Auswertung implementiert die in Kapitel 4 beschriebenen Erkennungsalgorithmen. Die Datenausgabe wiederum ist dafür zuständig die Ergebnisse im gewünschten Format auszugeben.

Der beschriebene Ansatz hat den Vorteil großer Flexibilität, besserer Übersichtlichkeit und Wartbarkeit. So können die Daten zwischen den einzelnen Stufen abgegriffen werden, um die Verarbeitungsschritte zu überprüfen. Dadurch, dass die einzelnen Stufen in separaten Prozessen laufen können die 4 Prozesskerne des Raspberry PI 2 außerdem optimal ausgenutzt werden.

## 5.3 Gesamtsystem

Der Sturzsensor ist Teil eines Gesamtsystems in welches dieser über WLAN eingebunden ist. Dabei kommuniziert der Sensor mit der Systemzentrale über Nachrichten welche in einem gemeinsam abgestimmten Format codiert sind. Bei einem erkannten Sturz wird eine kurze Meldung an die Systemzentrale geschickt welche dann die festgelegten weiteren Schritte ausführt.

Außerdem können von der Systemzentrale Konfigurationsnachrichten an den 3D-Sensor geschickt werden (z.B. Neustart, Übertragen von Messdaten, ...).

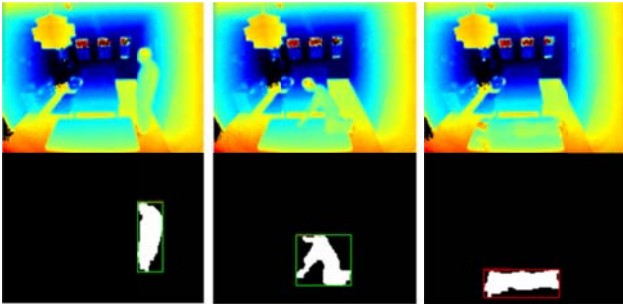
## 6 Experimentelle Ergebnisse

Um das entwickelte Sturzerkennungssystem zu konzipieren, optimieren und zu testen, wurden eine Vielzahl von verschiedenen Sturzscenarien mit dem Sensor aufgezeichnet. Das heißt, zahlreiche (gesunde) Personen haben verschiedene Stürze (z.B. Stolpern, zu Boden sinkend, ...) simuliert. Außerdem wurde auch in verschiedenen Räumen getestet, um zu verhindern, dass die Algorithmen auf bestimmte Situationen überangepasst werden.

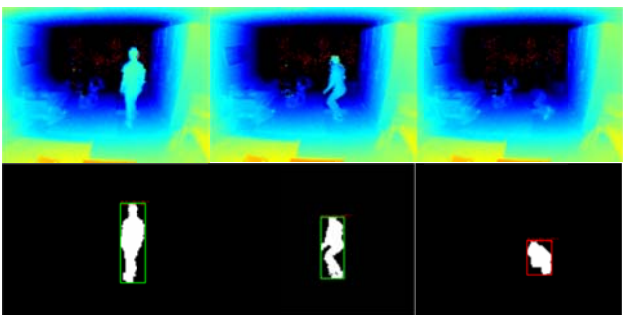
In **(Bild 6)** und **(Bild 7)** sind die Ergebnisse des Algorithmus für zwei verschiedene Sturzscenarien dargestellt. Wie zu sehen ist, kann die Person aus den Tiefendaten extrahiert, ihre Bewegung getrackt und ein Sturz erkannt



werden. Wie zu sehen ist dabei der Blickwinkel mit dem der Sensor den Sturz wahrnimmt nicht relevant.



**Bild 7** Von links nach rechts simulierter Sturz auf Matratze in einem Testraum. Oben die Daten der Tiefenkamera und unten das visuell dargestellte Ergebnis des Trackings und der Auswertung (roter Kasten um die Person zeigt einen Sturz an). In diesem Fall nimmt die Kamera den Sturz einer männlichen Person von der Seite wahr.



**Bild 8** Von links nach rechts simulierter Sturz auf Matratze in einem Testraum. Der Sensor nimmt den Sturz von hinten wahr.

## 7 Zusammenfassung

Erste Auswertungen der Ergebnisse zeigen, dass mit diesem System eine sehr gute Sturzerkennung möglich ist. Insbesondere konnten Personen bei ausreichender Sichtbarkeit zuverlässig getrackt und der Körpermittelpunkt sauber ermittelt werden. Bezgl. der Sturzklassifikation gilt es allerdings einige Abwägungen zu treffen. Setzt man den Grenzwert für den Körpermittelpunkt für einen Sturz sehr niedrig an, muss die Person nach einem Sturz sehr flach auf dem Boden liegen, sodass es als Sturz gewertet wird. Setzt man den Grenzwert höher an, können schnelle Bewegungen nach unten, z. B. um etwas unter dem Sofa hervorzuholen, einen Fehlalarm auslösen.

Natürlich gibt es auch Einschränkungen. Wenn z.B. die Sichtlinie des Sensors durch einen Gegenstand blockiert wird, z.B. durch einen Tisch, so kann das Geschehen dahinter nicht direkt erkannt werden.

Um die Qualität der Sturzerkennung besser beurteilen zu können, werden noch weitere Tests, auch in der Praxis, durchgeführt. Anhand dieser wird dann eine statistische Auswertung durchgeführt.

## 8 Literatur

- [1] Generali Altersstudie 2013: Wie ältere Menschen leben, denken und sich engagieren. Broschiert, 320 Seiten, Fischer Taschenbuch Verlag, Frankfurt 2012.
- [2] Pannurat, N.; Thiemjarus, S.; Nantajeewarawat, E. Automatic Fall Monitoring: A Review. *Sensors* 2014, 14, 12900-12936.
- [3] Dana Harry Ballard and Christopher M. Brown. 1982. *Computer Vision* (1st ed.). Prentice Hall Professional Technical Reference.
- [4] Gasparrini, S.; Cippitelli E.; Spinsante S.; Gambi E.; A Depth-Based Fall Detection System Using a Kinect Sensor. *Sensors* 2014, 14, 2756-2775.
- [5] Stone, E.E.; Skubic, M., "Fall Detection in Homes of Older Adults Using the Microsoft Kinect," in *Biomedical and Health Informatics, IEEE Journal of* , vol.19, no.1, pp.290-301, Jan. 2015
- [6] Xin Ma; Haibo Wang; Bingxia Xue; Mingang Zhou; Bing Ji; Yibin Li, "Depth-Based Human Fall Detection via Shape Features and Improved Extreme Learning Machine," in *Biomedical and Health Informatics, IEEE Journal of* , vol.18, no.6, pp.1915-1922, Nov. 2014