COMPUTER VISION-BASED FALL DETECTION METHODS USING THE KINECT CAMERA: A SURVEY

Salma Kammoun Jarraya

Department of Computer Science, King Abdelaziz University, Jeddah, Saudi Arabia

ABSTRACT

Disabled people can overcome their disabilities in carrying out daily tasks in many facilities [1]. However, they frequently report that they experience difficulty being independently mobile. And even if they can, they are likely to have some serious accidents such as falls. Furthermore, falls constitute the second leading cause of accidental or injury deaths after injuries of road traffic which call for efficient and practical/comfortable means to monitor physically disabled people in order to detect falls and react urgently. Computer vision (CV) is one of the computer sciences fields, and it is actively contributing in building smart applications by providing for image\video content "understanding." One of the main tasks of CV is detection and recognition. Detection and recognition applications are various and used for different purposes. One of these purposes is to help of the physically disabled people who use a cane as a mobility aid by detecting the fall. This paper surveys the most popular approaches that have been used in fall detection, the challenges related to developing fall detectors, the techniques that have been used with the Kinect in fall detection, best points of interest (joints) to be tracked and the well-known Kinect-Based Fall Datasets. Finally, recommendations and future works will be summarized.

KEYWORDS

Fall Detection, Kinect camera, Physically disabled people, Mobility aid systems

1. INTRODUCTION

Physical disability is defined as a significant and persistent physical condition that limits a person's movement[1]. Physical disabilities degree and type vary according to individual circumstances[2]. But it should be kept in mind, however, that persons with a disability may not be able to walk at all without the Ambulatory Assistive Devices (AAD)\mobility aids[3]. AAD such as crutches, canes, walkers, braces, and wheelchairs, are prescribed to persons for a variety of reasons: to decrease excessive weight bearing on the lower extremity, to correct the imbalance, to reduce fatigue, or to relieve pain secondary to the loading of damaged structures [2]. So, if the physical disables persons have something common, it will be using the mobility aid or what is called in physical therapy AAD because the reduced function of legs and feet. This paper focus on fall detection of physically disabled persons for two reasons[2]:

1. Physical disabilities are the most common disabilities (73%), followed by intellectual\psychiatric (17%) and sensory (10%). And that what makes author chooses the physical disability from different disabilities.

---

[1] The Oxford Dictionary of New Words.

[2] About Disability: https://www.health.wa.gov.au/publications/daip/training_package/fscommand/Disability.pdf

[3] Physical and mobility impairment factsheet. Available: https://www.google.com.sa/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjE0Iq56snQAhUELsAKHV8qA8oQFgggMAA&url=https%3A%2F%2Fwww.dorsetforyou.gov.uk%2Fmedia%2F169354%2FPhysical-and-mobility-impairment-factsheet%2Fpdf%2FPhysical_and_mobility_impairment_factsheet.pdf&usg=AFQjCNE_ZXSX3VqLrC9I29hleMQp_RaoEg

DOI: 10.5121/ijcsit.2018.10507

- The most people with disabilities experience some form of limitation or restriction due to their disability. Those people frequently report that they experience difficulty being independently mobile. And even if they can, they are likely to have some serious accidents such as falls. So, by the appropriate fall detection approach, the fear from fall can be prevented and gives the physical disable person the independent to mobile in his/her room, office or apartment without the help of another person to stay and maintain him\his. Furthermore, in this paper, to be more specific, the target is physically disabled users who use a cane as a mobility aid, also for two reasons: (i) Canes are the most commonly used mobility aids [3]. (ii) Physical disability occurs frequently in older adults [4] who use the cane regularly, and have the highest risk of death or serious injury arising from a fall; and that risk increases with age [3].

- Disabled people, in fact, can overcome their disabilities in carrying out daily tasks in many facilities [1]. However, for some emergencies like fall which is one of the most frequent emergency, a disabled person requires urgent assistance without an explicitly ask for help. Indeed, according to 2017 statistics from the World Health Organization (WHO)[4], one out of three 65-year-old people falls each year and, as age increases to 80, the fall occurs each year. And with increases of age, the frequency of falls increases. Furthermore, falls constitute the second leading cause of accidental or injury deaths after injuries of road traffic. These statistics call for efficient and practical\comfortable means to monitor physically disabled people in order to detect falls and react urgently.

The most popular Fall Detection (FD) approaches are classified based on the sensor used, and they organize in three categories: wearable device based approaches, ambient sensor-based approaches, and camera (vision) based approaches [5]. And because the highly obtrusive of the wearable sensors [6] and the uncomfortable during normal daily life activities; and the highly sensing of the ambient sensors which sense the pressure of everything in and around the person and generating too much false alarms [5], this paper considers only camera sensor approaches and more specifically the Kinect camera for its advantages which will discussed later.

This survey may help researchers to answer to the following research questions:

1. What are the most popular approaches that have been used in fall detection?
2. What are the challenges related to developing fall detectors?
3. What are the techniques that have been used with the Kinect in fall detection?
4. What are the best points of interest (joints) to be tracked?
5. What are the well-known Kinect-Based Fall Datasets?

This survey consists of this introductory section in addition to four further sections. Section 2 (Background) introduces, in general, a brief background about the most popular fall detection approaches based on the sensor used; and focused on the Kinect camera as one of the vision sensors. The Kinect hardware and software, and the skeleton tracking using Kinect will be presented. Also, presents author reasons to the reader that make researchers choose the Kinect camera and its official SDK to detect the fall of the physically-disabled cane users.

Section 3 (Literature Review on Kinect-Based Fall Detection) overviews the related works on FD that used Kinect; their limitation\disadvantages. Followed by a discussion of the existing Kinect-based fall datasets. Finally, the skeleton joints and features that used for FD are overviewed. In Section 4 (Discussion and Conclusion), discussion, recommendations and future works will be summarized.

[4] World Health Organization (WHO): Falls Statistics: http://www.who.int/mediacentre/factsheets/fs344/en/

2. BACKGROUND

This section aims to introduce the background that the readers need to know about the fall detection approaches in general and the Kinect camera as acquisition sensor. Then it presents the reasons that lead researchers to choose this camera and its Software Development Kit (SDK) from other cameras and libraries to track people with a physical disability to detect fall.

More specifically, this section is organized as the following: Section 2.1 overview the most popular approaches used to detect the fall based on the sensor used. Also, presents the different fall scenarios in Section 2.1.1, the challenges related to FD systems in Section 2.1.2, and explain our reasons to the reader that make author focusing on the Kinect sensor from all other sensors that have been used to detect the fall (Section 2.1.3). While Section 2.2 first introduces the Kinect camera, its components and its two generations (Section 2.2.1). Section 2.2.2 contains the different libraries used with Kinect and focus mainly on Kinect's official SDK. And at the same section, the skeleton tracking and how it varies when using different Kinect versions or different Kinect libraries are overviewed.

2.1. THE STATE OF ART ON FALL DETECTION

FD system is an assistive device whose primary objective is to alert when a fall event has occurred. Several studies reviewed existing FD systems. Mubashir et al.[5], Igual et al. [7], Zhang et al. [8], and Noury et al. [9] surveyed FD systems using different review criteria for fall types and used devices. Mubashir et al. [5], introduced different types of falls, including falls from walking, standing, standing on supports (*e.g*., ladders), sleeping, lying in bed, or sitting on a chair. In addition, they classified FD approaches into three categories based on the used device (see Figure 1): wearable device-based, ambience sensor-based, and camera (vision) based approaches. Igual et al. [7] identified three other fall types (forward, backward and sideways), and they classified the reviewed FD systems into two types: context-aware systems and wearable devices. Zhang et al. [8] classified the reviewed FD methods into either non-vision sensor-based, or exclusively vision sensor-based. They reviewed vision-based methods only which they classified into three categories based on their camera type: single RGB camera-based, 3D-based using multiple RGB cameras, and 3D-based using depth cameras. In contrast, Noury et al. [9] classified the reviewed FD methods in terms of their technique, either analytical or machine learning methods. In addition, they obtain 20 fall scenarios\types from the literature researchers. Their scenarios/types were categorized as: backward fall (both legs straight or with knee flexion), forward fall, lateral fall to the right, lateral fall to the left, syncope and neutral.

Based on the above studies [5, 7-9], the advantages and disadvantages of existing FD approaches which classified into three categories depending on the used sensor are summarized in Figure 2. Based on this summary, this study focus on the third class of approaches (vision-based) and, in particular, the Kinect skeleton-based approach. Indeed, ambient sensor-based approaches sense the pressure of everything in and around the person and generating too much false alarms [5]. As for wearable sensor-based approaches, they may have limited acceptance by the users for three main reasons: the highly obtrusive nature of wearable sensors, their incurred discomfort during normal, daily life activities [6], and their numerous false alarms. In contrast, a Kinect skeleton-based approach (in the third category) offers multiple advantages. So, in this paper considers only a Kinect as acquisition sensor for its advantages which this paper will focus on and discuss later in this section after presented fall scenarios and the challenges related to fall detection.

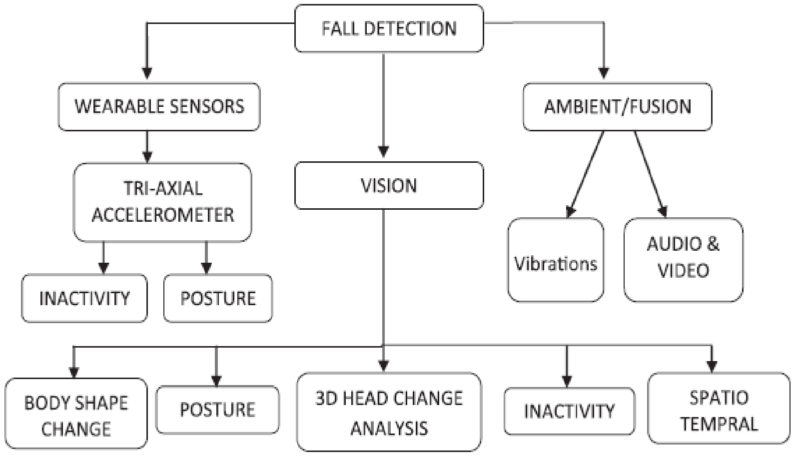


Figure 1: Classification of fall detection methods from [5]

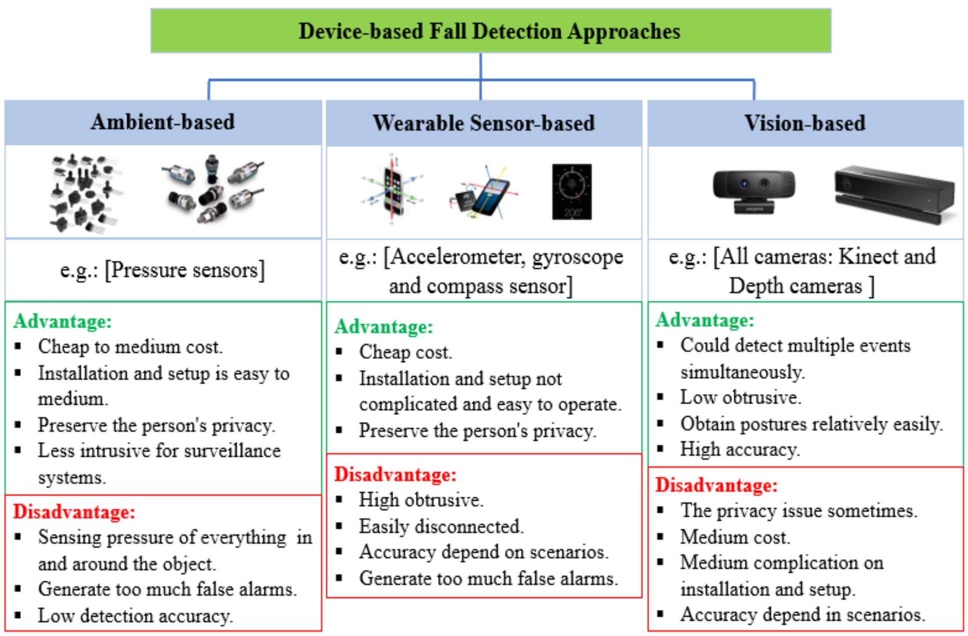


Figure 2: The current fall detection methods classification into three device-based approaches: ambient-based, wearable sensor-based, and vision-based approaches

2.1.1 Fall Scenarios

The fall scenarios are numerous, but they share common characteristics. Some of them represent true fall situations (positive situations) whereas others represent ‘pseudo’ fall situations (negative situations) [10]. In addition, some fall characteristics also exist in normal actions (negative situations), e.g., a crouch also demonstrates a rapid downward motion [7]. Table 1 shows some fall and non-fall scenarios that are collected from the literature [2, 5-11, 7, 8]. Non-fall contains some daily activity when changing from one posture to another. ADL contains Activity of Daily Living that is not specified by the authors.

Table 1: Fall and Non-Fall Activities\Scenarios from the Literature

(✓: The paper mentions\detects this scenario, - : The paper doesn't mention this scenario, ×: The approach in the paper didn't use or can't detect this scenario)

References **Scenarios:**		**[5]**	**[7]**	**[11]**	**[12]**	**[13]**	**[14]**
Falls	Falls from walking or standing	✓	-	✓	✓	✓	✓
	Falls from standing on support	✓	-	-	-	×	-
	Falls from sitting on a chair	✓	-	×	✓	✓	✓
	Falls from sleeping or lying	✓	-	✓	-	✓	-
Non-Falls	From walking or standing to sitting	-	-	-	✓	-	✓
	From sitting to standing	-	-	-	-	-	-
	From sitting to lying (include from walking or standing to lying because s/he should sit first)	-	-	-	✓	-	-
	From lying to sitting (include lying to walking also, because if person lying and want to walk, s/he should sit first)	-	-	-	-	-	-
	Sitting on the floor – both legs folded behind	-	-	✓	✓	-	-
	Kneeling on the floor (crouching down)	-	-	✓	✓	-	-
	Squatting	-	-	✓	-	-	-
	Bending down (to wear shoes or tie shoelaces) or picking up an object from the floor	-	-	✓	✓	-	-
ADL	Activity of Daily Living	-	-	✓	✓	✓	✓

2.1.2 Challenges Related To Fall Detection

Many challenges [1,13, 7, 15]should be taken into consideration when developing a robust FD solution. These challenges are:

1. Performance: FD system must be as accurate and reliable as possible. A robust fall detector should show both high sensitivity and specificity
2. Usability: fall detector should be with no restrictions regarding the room or the user position. And in term of skeletal tracking initialization, it should be automatic and do not require users to enter a calibration pose.
3. Acceptance: the acceptability, practicality, and comfortability of the technology should take into consideration. Vision systems as other non-intrusive methods are excellent in this sense.
4. Privacy: in healthcare application, the user privacy is a significant issue. So, the fall detector must be developing a solution for protecting privacy. Using Kinect, this could be solved by using only depth and skeleton information.

5. Security: different levels of security should be decided to different users in such complex systems. So, the communication links and the infrared used by depth camera should be made secure and more reliable.
6. Flexibility: fall detector must support the use by new users without the need of re-training the system.

2.1.3 WHY THE KINECT

The Kinect camera is a human motion tracking peripheral which was first developing for the Xbox 360 console [16]. It provides full-body 3D motion capture, facial recognition, and voice recognition capabilities [17]. It has been used in various systems, and the Kinect has been yet the first camera which combines the RGB color camera with structure-light depth camera into one single camera [18].

A Kinect is a sensor as other sensor have its advantages and disadvantages[11,12,13,16,18-21] **.** Many authors and programmer make the Kinect camera their first choice because it has been yet the first camera which combines all the following advantages:

- Kinect combines multiple streams: since the Kinect camera combines the RGB color, depth, skeleton, infrared, body index, and audio into one single camera [18], make the Kinect the first available camera in all these streams.
- Kinect give RGB-D images: containing the depth information and the visual RGB information which takes the advantages of the RGB image that provides appearance information of an object and also the depth image that is immune to the variations in color, illumination, rotation angle and scale [18]. Combining RGB and depth information can dramatically improve the classification accuracy [19] which will help the developers to improve the performance of their system.
- Suitable to the indoor application: with a comparison to another depth camera, it was found that the depth technique in Kinect is more suitable to the indoor application [18]. And that what makes the Kinect-based fall detection system suitable to be used in hospital and the care homes.
- Safe sensor: the infrared (IR) laser in the Kinect uses a type 2 infrared light which is safe to use on the human, not like most of the laser scanners which use the type 1 laser that is dangerous to eyes when no protecting has been worn [18]. And this will ensure the safety of the user from any system using Kinect as the acquisition sensor.
- Low obtrusive: all the camera including Kinect are low obtrusive not like the wearable sensors which can be highly obtrusive [6] and uncomfortable during normal daily life activities. And by using the Kinect, researcher will ensure the comfortability to the physically disabled persons who already not comfortable with their disability.
- Availability and the low cost: the Kinect is widely available and has a small price which is 150$ including the adapter [17, 18]. And with the free proposed software, it makes this meager cost for a complete hardware and software FD system.
- Preserves the person's privacy: if only depth images are used, it protects the person's privacy by producing a stickman display or even a grey image [12].
- Independent of external light conditions: using the infrared light, the Kinect sensor is capable of extracting the depth maps in dark rooms [12]. So, it can detect falls in the late evening or even in the nighttime.

- Portable: Other reasons for choosing Kinect because it is portability and the easiest in hold it and change its place [18].
- Fast frame-rate: it could use for real-time because of the fast frame-rate (FPS at 30Hz) [12].
- Reasonable range: from 0.4m to 4.5m [16, 18]. Which suit to use in bedrooms and home care rooms.
- Accurate depth information 1mm error after calibration: after calibration, the re-projection error is 0.16 pixels and 0.17 pixels. These results show that the lenses, which are used in Kinect, are already perfect since a typical webcam has a re-projection error greater than 0.7 pixels.

Next section reviews extensively the hardware and software of the Kinect, and how this camera does the skeleton tracking.

2.2. KINECT CAMERA

This section overviews in detail, the Kinect hardware\components and its two generations, the Kinect software\libraries and how the skeleton is tracking when using different Kinect software.

2.2.1 KINECT HARDWARE

The Kinect refer to the camera which has been provided by Microsoft in 2010 for the Xbox 360 console, and it is known as Kinect v1 (Figure 3.a). In 2014, Microsoft developed the Kinect v2 (Figure 3.b) which improve some hardware and software features [16]. Kinect hardware contains RGB camera, IR camera, IR projector, multi-array microphone and a tilt moto [16, 19, 22-24]:

- RGB camera: provide colorful images with resolution 640×480 pixels for v1 at 30 Hz. And colorful pictures with resolution 1920×1080 pixels for v2. Both contain 30 frames\s.
- IR camera and IR projector: together they provide the 3D depth sensor, which offers depth images with resolution 320×240 pixels for v1 and has a ranging limit of 0.8 ~ 3.5 m and has the Field of View (FOV) 62º(h) x 48.6º(v). And for v2, it provides depth images with resolution 512×424 pixels and has ranging limit of 0.4 ~ 4.5m. Both have the frame rate of 30 frames\s. The v2 increased FOV of 70º(h) x 60º(v).
- Multi-array microphone: consist of four microphones to make live communication available. It can process up to 16-bit audio signals at 16 kHz sample rate.
- Tilt motor: it is existing only on Kinect v1 to allow the sensor adjustment and can tilt the sensor up to 27º either up or down.

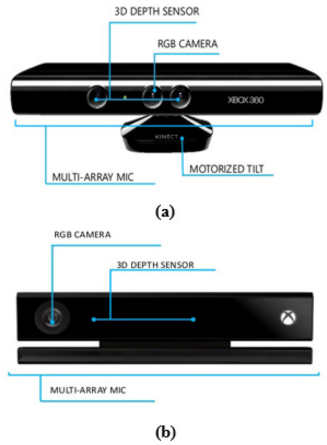


Figure 3: The Kinect cameras: (a) The Kinect v1 camera[5], and (b) The Kinect v2 camera[6]

The above hardware features make the Kinect able to provide full-body 3D motion capture, facial recognition, and voice recognition capabilities [17]. Which make the Kinect very useful in various application such as: object tracking and recognition [19, 21, 22], human skeleton tracking and activity analysis [16, 21, 25], hand gesture analysis [16, 26], 3D-simultaneous localization and mapping [19, 22], emergences detection such as: assault detection [27, 28] and fall detection [11, 13, 29], and other.

And so far, there have been only two generations of the Kinect, one for the Xbox One (Kinect v1) and one for Windows (Kinect v2), the first used on a gaming device, and the second used on the Windows machines. But in 2015, Microsoft announced that they will stop producing separate versions of the Kinect but will instead encourage developers to purchase the Kinect for Windows adapter instead to plug their Kinect into a PC . For the Kinect v2 with the Windows adapter, it's recommended to use the Kinect for Windows SDK 2.0 . The differences between the two generations of the Kinect camera can be shown in Table 2 [16, 19, 22-24].

2.2.2. KINECT SOFTWARE AND SKELETON TRACKING

For the Kinect software, after the Microsoft has released the Xbox360, some companies provide unofficial free libraries and SDK for the Kinect such as CL NUI Platform, OpenKinect\Libfreenect, OpenNI, and PCL. In 2011, the Microsoft released the official Kinect SDK [19]. The Kinect SDK 2.0 enables developers to create applications that support voice and gesture recognition, using Kinect sensor technology on computers running Windows 8, Windows 8.1, and Windows Embedded Standard 8. It also includes Application Programming Interfaces (APIs), device interfaces and code samples. It also provides free tools which can be used for detecting and tracking the body\skeleton and the head of a person [16]. For more about Kinect free libraries see [19].

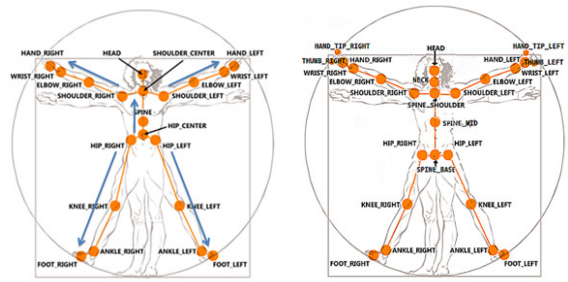
[5] Introduction to Kinect - Update v 1.8: https://www.slideshare.net/MatteoValoriani/introduction-to-kinect-update-v-18

[6] Develop store apps with Kinect for windows: https://www.slideshare.net/MatteoValoriani/develop-store-apps-with-kinect-forwindowsv2-150601152707lva1app6891

Table 2: Comparison Between the Two Kinect Camera Generations, Kinect v1, and Kinect v2

Kinect Generations		
Properties	**Kinect v1**	**Kinect v2**
Shape		
Frames Per Second (fps)	30fps	30fps
Color Resolution	640 × 480	1920 × 1080
Depth Resolution	320×240	512 × 424
Sensor	Structured light	Time of flight
Range	0.8 ~ 3.5m	0.5 ~ 4.5m
Field of View	62º(h) x 48.6º(v)	70º(h) x 60º(v).
Audio Streams	4-mic array 16 kHz	4-mic array 48 kHz
Tilt Motor	Motorized	Manual
Number of Apps	Single	Multiple
Body Tracking	2 people	6 people
Body Index	6 people	6 people
Joints	20 joints per people	25 joints per people
Hand State	Open, closed	Open, closed, Lasso
Aspect Ratio	4:3	6:5
Supported OS	Win 7, Win 8	Win 8
USB Standard	2.0	3.0

Fall detection approach as most of the computer vision applications for Human Activity Recognition (HAR) which recognize human activities through skeleton tracking by representing body parts as joints. They exploit the fact that body features calculated from a 3D skeleton increase robustness across persons and can lead to higher performance [6]. Skeleton tracking can be offered from the previous different libraries such as unofficial OpenNI (Open Natural Interaction) and official Kinect SDK and different Kinect versions such as v1 and v2 [19]. But each has its advantages and disadvantages in skeleton tracking. Kinect v1 can recognize up to six users and tracks up to two users in details [16] and detect 20 joints per people. While Kinect v2, can recognize and tracks up to six users in details and detect 25 joints per people [19]. For the Kinect libraries, Kinect SDK 1.8 and older track the full body include head, hands, feet, clavicles and can calculate 3D positions of 20 joints per people (Figure 4.a) [16]. While SDK 2.0, do the same but with 25 joints per people (Figure 4.b). On another hand, OpenNI does not support the track of head, hands, feet, clavicles but support the hands only mode [19] and calculate 3D positions and rotation of 15 joints per people [21]. Table 3 shows the difference between the two versions of Kinect SDK 1.8 and older, and 2.0 in skeleton tracking.



(a) (b)

Figure 4: The positions of the skeleton joints of the user's body in the camera's FOV: (a) The positions of the 20 joints in Kinect SDK 1.8 and older[7], and (b) The positions of the 25 joints in Kinect SDK 2.0[8]

Table 3: Comparison Between the Kinect SDK Versions

Versions	**SDK 1.5, 1.6, 1.7, and 1.8[10]**	**SDK 2.0[11]**
# of Joints	20	25
Tracked Joints	The supported joints are the following: • Shoulder-Center • Spine • Head • Hip-Center • Shoulder-Left • Elbow-Left • Wrist-Left • Hand-Left • Shoulder-Right • Elbow-Right • Wrist-Right • Hand-Right • Hip-Left • Knee-Left • Ankle-Left • Foot-Left • Hip-Right • Knee-Right • Ankle-Right • Foot-Right	The supported joints are the following: • Spine-Base • Spine-Mid • Neck • Head • Shoulder-Left • Elbow-Left • Wrist-Left • Hand-Left • Shoulder-Right • Elbow-Right • Wrist-Right • Hand-Right • Hip-Left • Knee-Left • Ankle-Left • Foot-Left • Hip-Right • Knee-Right • Ankle-Right • Foot-Right • Spine-Shoulder • Hand-Tip-Left • Thumb-Left • Hand-Tip-Right • Thumb-Right
Compatible with	Kinect v1	Kinect v2

---

[7] Tracking Users with Kinect Skeletal Tracking: https://msdn.microsoft.com/en-us/library/jj131025.aspx

[8] JointType Enumeration: https://msdn.microsoft.com/en-us/library/microsoft.kinect.jointtype.aspx?f=255&MSPPError=-2147217396

The Kinect official SDK 2.0 is recommended to track the users' skeletons for several reasons:

1. Kinect official SDK is more advantageous than other libraries when requiring skeleton tracking [19]. It provides full body tracking include head, hands, feet, clavicles. It detects 25 joints per people [16] which give more flexibility and options when you develop FD approach.
2. It seems to be more stable regarding collection quality of the original image and the technology for pre-processing [19]. And the quality images lead to better performance.
3. It can process some details such as occluded joints [19] meticulously. And that will be so useful when some joints are occluded or not seen by the Kinect.
4. It can use to producing a stickman display rather than the person itself [11] and that very useful for maintaining user's privacy.
5. Multiple Kinect sensors can be supported [19]. And this is a useful advantage on the system that needs another Kinect camera.
6. It doesn't need specific pose or calibration action to be taken for a user to be tracked [13]. Not like other using tracker which need to start the activity with a surrender pose [21] which will be difficult for our users because of their health conditional.

Next section focuses on the literature reviews on Kinect-Based fall detector and datasets. Also discusses different joints and features of the skeleton that have been used to detect the fall.

3. LITERATURE REVIEW ON KINECT-BASED FALL DETECTION

This section mainly focused on computer vision-based FD methods using the Kinect camera, review and discuss the FD datasets that recorded by the Kinect camera. Also presents the used skeleton joints and features for FD.

The literature review on Kinect-based FD approaches covers works in the years range to be from 2013 until now because, until it, they had FD accuracy rates of 70% to 80% only [11]. More specifically, this section presenting the related works on three essential aspects. First, Section 3.1 shows works on FD related to Kinect. Second, Section 3.2 overviews and discusses existing Kinect-based fall datasets. Then, Section 3.3 discusses the skeleton joints and features that used for FD.

3.1. KINECT-BASED FALL DETECTION APPROACHES

Main-body text is to write in fully (left and right) justified 11 pt. Times New Roman font with a 6pt. (paragraph) line spacing following the last line of each paragraph, but a 12pt. (paragraph) line spacing following the last paragraph. Do not indent paragraphs.

Several recent well-known studies [29-32] classified under vision-based approaches, use the Kinect for developing FD systems. This section reviews and analyzes those related Kinect-based fall detection methods. And then summarizes their disadvantages and limitations and proposed some solution to solve it.

C. Lee and V. Lee [11] present a system to detect falls and notify healthcare services or the victim's caregivers to provide help. They use the Kinect camera v1 as the input sensor and Microsoft Kinect SDK that provides their FD system with skeleton data. They choose the hip center joint to be tracked and processed the hip data received by two functions-based centers of mass. The first function checking if the hip position\coordinates within a certain threshold distance from the floor. And the second separates the user's hip center velocity into two components: vertical and horizontal velocities and then checks if the combined scores for both the vertical and horizontal components exceed their pre-set overall threshold score. If both functions return a fall, then a fall is tentatively indicated. After that, their postural recognition algorithm is then applied to reduce the number of false positives (such as sitting on the floor with both legs folded behind, kneeling on the floor, squatting, bending down to wear shoes or tie shoelaces) returned by their FD algorithm. For this purpose, two main defining features were identified and checked against previously obtained skeletal data of postures to ensure their specificity. The first feature, user's Ankles have to be below the user's hip center. The second feature, one of the user's legs is either folded below his body, or one knee is significantly higher than the other. Their method detect falls 100% in certain situations (fall from a bed or in open space) with specificity rate of up to 90%, that means their overall FD accuracy rate is 95%.

They concern to detect the three kinds of falls that frequently occur within the Home. These three kinds of falls are: falling from the bed, falling in open spaces (which their system detect), or falling off the chair and that fall is undetected by them due to the fall taking place in a low position; and reduced time required by the system to measure a significant increase in user's velocity while he is falling, hence resulting in the system not being able to detect the fall correctly. Falls off the chair is an essential fall scenario and one of our concerns for people who have a physical disability and using a chair most of the time for resting.

In 2014, in an effort to reduce the number of false alarms by collecting more information, Kwolek and Kepski [12] added to the Kinect a wearable smart device containing accelerometer and gyroscope sensors; this intelligent device is worn near the pelvis region of the monitored person. They use a triaxial accelerometer to indicate both a potential fall and whether the person is in motion, and the Kinect camera v1 with OpenNI library acquire depth images to reduce the number of false alarms and employ it whenever it is only possible. Their FD system runs under Linux OS. First, the data acquisition from the wearable device and transmitted wirelessly via Bluetooth to the processing device; and the depth acquisition from the Kinect which is connected via USB to this device. Then median filtered the depth image to fill the holes and smooth it. After that, store it in the circular buffer for further feature extraction and continuously updates it. While the person in motion, the system will continue extracting the foreground through subtraction to determines the connected components\objects. If the scene changes, the depth reference image will update. Otherwise, a potential of fall is examined using accelerometric data. If the measured acceleration is higher than the assumed threshold value equal to 3g, then recognize potential of fall and remove all the connected components in a binary image except the largest one which represented the segmented person. Then from the depth image, extract the correspond v-disparity image to calculate the floor plane parameters. After that, the system calculates some depth features, and classify the image\frame using SVM classifier to trigger the fall alarm.

Their algorithm achieves 98.33% accuracy, 100% sensitivity, and 96.67% specificity when using accelerometer and depth data, and 90% accuracy, 100% sensitivity and 80% specificity when using depth only which is the worst result compared to other techniques in their research. The 98.33% accuracy rate was obtained not by using the Kinect alone but also with the help of accelerometer and gyroscope sensors which is not acceptable because of the need to wear and carry various uncomfortable devices during normal daily life activities, the elderly may forget to wear such devices or determine not to be worn during the sleep which lack the ability of such detectors to detect a fall. In addition, they use foreground which needs to determine whether a foreground object actually is or is not a person, and in some cases, is or is not the person under the study. And when using the Kinect only, they achieved 80% specificity which means they could not avoid the false alarm by 20% (false alarm rate = 0.2).

Le and Morel [33] present a novel FD system based on the Kinect v1 sensor. First, they compute the room floor plane using the Kinect's floor plane equation. Second, they extract the head and spine joints coordinates and convert them to floor coordinates. Then, they calculate the distance from the floor and velocity features. After that, they classify the frames using the SVM. From their results, they first obtained 83.56% accuracy, 91.12% sensitivity, and 76% specificity. And after they remove the fall-like samples (sit and lie down the ground) from the training data, their results get higher to 91% accuracy, 100% sensitivity, and 82% specificity. Finally, they detect a fall on a duration of time to solve the misclassification of "Lie down the floor" which frequently classified as a fall activity, and they get 98.35% accuracy, 100% sensitivity, and 96.7% specificity.

The disadvantage of their work was the dataset that used for algorithm training which contains only nine scenarios: four falls, two falls-like, and three ADL (Activity of Daily Living). Forthermore, their scenarios performed by their subjects in a very specific way by marketing the ground using five points as shown in Figure 5; and in each scenario, the subject must follow these points and fall in a specific direction which make it less natural. That means, even when changing of the subject, the scenario will be performed at the same exact way.

N^0	Activity	Description
1	FallFront	S–>A–>Fall in the direction D
2	FallBack	S–>A–>Fall before arriving at D in your back in the direction of A
3	FallRight	S–>A–>Fall right after A in the direction C
4	FallLeft	S–>A–>Fall before arriving at C in your back in the direction of A
5	Sit	S–>A–>C–>Sit down suddenly on the floor
6	Lie	S–>A–>C–>Lie down in the direction of A
7	Walk	S–>A–>C–>A and stay at A position
8	Pick	S–>A Move to pick up an object, rise and stay at A
9	Sit	S–>A–>B–> Sit down and stay

Figure 5: Experimental environment in work [33] with the five markers (left) and the performed scenarios (right)

In 2015, Stone and Skubic [13] developed a two-stage FD system for detecting falls using the Kinect v1 in the homes of older adults. The first stage characterizes the vertical state of a 3D object for an individual frame using three features: the maximum height of the object, the height of the object's centroid, and the number of elements of the discretized (floor) plane. Then, using the vertical state of the tracked object over time series, it segments on ground events after filtering the vertical state time series using median and average filters. The second stage extracted five features from an on-ground event to generate confidence that a fall preceded it. Those features are minimum vertical velocity, maximum vertical acceleration, mean vertical velocity, occlusion adjusted change, and minimum frame-to-frame vertical velocity. Finally, using those five features and an ensemble of decision trees, fall confidence is computed for each on the ground event. As a preprocessing step, this system segments 3D foreground objects from each depth frame using dynamic background subtraction. When the falls are near the sensor and not significantly occluded, this system can achieve 98%, 70%, and 71% cross-validation accuracy detection of standing, sitting, and lying falls, respectively; however, when the falls are far to the sensor and significantly occluded, the system can achieve 79%, 58%, and 5% cross-validation accuracy detection of standing, sitting, and lying falls, respectively. They reached a shallow false alarm (only one per month) but the FD accuracy is not high enough especially for sitting and lying positions, and when the faller is far from the Kinect. The low accuracy when the faller is far because of the 320×240 depth resolution used in this system to keep the space required to store the data to a manageable. And the low accuracy detection of sitting because the presence of the chair used by the faller which is often part of or the entire chair will be identified as foreground; and that will result the fall. Finally, the low accuracy detection of sitting because of the significantly reduced fall motion.

Existing FD systems using Kinect [11-13, 33] differ in their performance. Some do not cover many fall types and scenarios; others have high false alarm rates when operating on particular postures, yet others have low accuracy when the faller is far from the Kinect, or even trained with the less natural dataset. In addition, those trying to improve the FD rate use wearable sensors, which may hinder daily activities and make the person uncomfortable. Furthermore, the proposed systems' performance evidently depends on the features used. However, the feature differences (in nature and number) and the lack of a benchmark dataset hinders a systematic evaluation of the performance of existing systems.

FD with the Kinect is a complicated process for which presently there is no an identical solution. As such, a FD method must: suits physically disabled people, relies solely on Kinect, and can determine various types of falls in different postures with a low false alarm and a high accuracy. Author suggests in Table 4 some solutions to avoid the preceding disadvantages of the literature studies as shown.

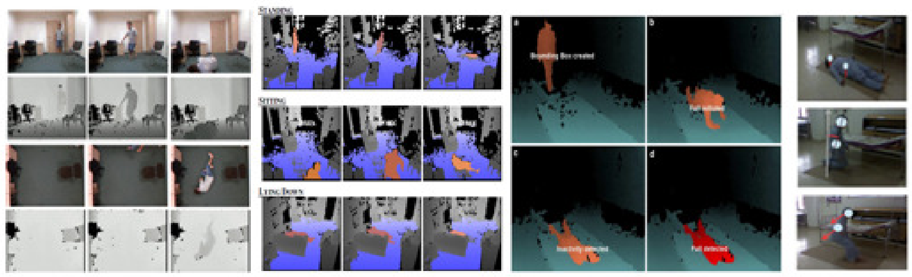
Table 4: Disadvantages and Limitations of the Previous Fall Detection Approaches

Reference	Sensors + libraries	Disadvantages and Limitations	Solution
C. Lee and V. Lee [11] (2013)	Kinect camera v1 + Kinect SDK	Cannot detect when user falls off the chair	Increase the points of interest (track more joints from skeleton rather than only the hip joint) and use another feature rather than the distance from the floor
Kwolek and Kepski [12] (2014)	Kinect camera v1, Triaxial accelerometer, and gyroscope + OpenNI	Using many sensors included two wearable sensors (which is uncomfortable devices and cost a lot)	Using the Kinect only (which will decrease the cost and unobtrusive)
		When using the Kinect only, they could not avoid the false alarm by 20%	Using skeleton stream rather than depth which is better on person tracking and will reduce the false alarm rate
		The foreground needs to determine whether a foreground object is or isn't a person, and in some cases, is or is not the person under the study	Use skeleton tracking from SDK (which track only the person and could determine the person under the study by using the tracked id) than foreground technique
Le and Morel [33] (2014)	Kinect camera v1 + Kinect SDK	To solve the misclassification of fall-like which frequently classified as fall activity, they have to detect a fall in a duration of time which increases the computations	Because they use the distance from the floor features, their system will consider the fall-like as fall. Investigate different features that do not detect the fall-like as fall
		The dataset scenarios perform in particular way which makes it less natural	Using more natural dataset by allow the subject to perform the scenario on his adaptive way
Stone and Skubic [13] (2015)	Kinect camera v1 + libfreenect library	Low accuracy when the faller is far because of the Kinect with 320×240 depth resolution used	Using the newest Kinect version (v2) which has 512×424 pixels depth resolution
		Low accuracy detection of sitting because the presence of the chair used by the faller which is often part of or the entire chair will be identified as foreground which results in the fall	Use skeleton tracking from SDK which tracks only the user rather than foreground technique which identify some objects as part of the user
		The low accuracy detection of lying because the significantly reduced fall motion	Avoid it by skeleton tracking which only interest on the positions of the skeletal joints

3.2. KINECT-BASED FALL DETECTION DATASETS

Since 2011, there have been different datasets which were taken by the Kinect camera. Cai et al. [19] surveyed 46 existing RGB-D benchmark datasets among which 20 are well elaborated. These RGB-D datasets contain the depth information and the visual (RGB) information, and the Kinect sensor can acquire them. They have divided it into five categories depending on the facilitated computer vision applications. These categories include object detection and tracking, human activity analysis, object and scene recognition, SLAM (Simultaneous Localization and Mapping) and hand gesture analysis. In human activity analysis category, Cai et al. in [19] overviewed the UR Fall Detection dataset created by the University of Rzeszow in 2014 [21]. This dataset was created by Kwolek and Kepski to detect and recognize the human fall. It was recorded using two Kinect cameras and corresponding accelerometric data in rooms such offices, classrooms, etc.The first camera was parallel to the faller within a 1m distance

from the floor, and the second camera was on the room's roof within a 2.5m distance from the floor to cover all the room. This dataset contains 70 sequences: 30 sequences represent the fall situations which was recorded by the two cameras, and 40 sequences represent the ADL which recorded by only one camera. This dataset was performed by five healthy subjects who acted two types of fall: fall from standing and fall from sitting on a chair. Some images from this dataset are shown in Figure 6.a.



(a) (b) (c) (d)

Figure 6: Samples of previous Kinect-based fall detection datasets: (a) Image samples of RGB and depth images from UR fall detection dataset [12], (b) Image samples of depth images performed in apartments of older adults from [13], (c) Image samples of infrared images from the dataset in [29], and (d) Image samples of fall scenario from dataset in [33]

C. Lee and V. Lee [11] used a preliminary dataset to optimize their enhanced FD algorithm. This dataset includes 34 fall and non-fall events of 10,479 frames skeletal data captured by Kinect.

Stone et al. [13] present a method for detecting falls in the homes of older adults using the Microsoft Kinect and a two-stage FD system. They deployed a Kinect camera and a computer in the 13 elderly resident's homes at an independent living facility. The 16 residents' ages were from 67 to 97. Seven were male, and nine were female. The 3,339 days of continuous data containing 454 (near and far) falls include 14 standing, five sitting, and two lying fall scenarios. Also, it contains three scenarios of fall-like. The Kinect is placed above the front door on a small shelf a few inches below the ceiling in distance height 2.75 m from the floor, and the computer placed above the refrigerator. The authors use only the values from the Kinect depth stream through the open source libfreenect library. Some examples of the depth images shown in Figure 6.b.

Mastorakis and Makris [29] attached the Kinect to a tripod at a 2.04m distance from the floor and at no farther than 7m from the area of possible fall. They capture 184 videos include: 48 falls, 12 slow falls, 48 lying activities on the floor, 32 seating activities, 12 slow activities and 32 picking up an item from the floor. The fall types include: forward, backward and sideways. Eight subjects performed their dataset, two of them performed in slow motion to simulate the behavior of an older adult. Their data was in infrared streams. An example of this dataset shown in Figure 6.c.

Le and Morel [33] present a method to detect fall using skeleton data and SVM technique. They evaluate their proposed method using dataset recorded using Kinect v1. Their dataset contains nine activities (four falls, two fall-like, and three ADL) performed in particular ways. Fall activities include back, front, right, and left falls; and fall-like includes sit and lie down actions; while ADL activities include walk, sit down the bed, pick an object. The nine activities\scenarios performed twice by six subjects with 20 to 35 ages in a 3x3m room with a bed; the room marked by five points on the ground, so the subjects can follow when performing the scenarios. Finally, they end with fall dataset contain 108 videos. An example is shown in Figure 6.d.

Overall, there were different fall datasets recorded using the latest technology Kinect [11-13, 29, 33]. As summarized in Table 5, they differ regarding their coverage, size stream types, and availability. This table also highlights a need for a fall dataset that is available to researchers, that suit the cane users, that covers many fall types and scenarios, and most importantly, that recorded in all the streams provided by the Kinect camera to provide for the needs of different techniques.

Table 5: Comparison between the Existing Kinect-Based Fall Datasets

Dataset	[12]	[11]	[13]	[29]	[33]
Availability	Yes[9]	No	No	No	No
Coverage	Falls from standing and sitting	-	Many fall types	Many fall types	Many fall types
Size	70 videos	10,479 frames	Videos of 3,339 days	184 videos	108 videos
Streams	RGB and Depth	Skeleton	Depth	Infrared	Skeleton
Special Type	Healthy people	Elderly	Elderly	Healthy people and Elderly	Healthy people
Setting	Artificial	-	Real word	Artificial	Artificial

3.3. FALL DETECTION USING SKELETON JOINTS AND FEATURES

The success in human activity recognition remains dependent on the correspondence between the human activities and the used joints features. Indeed, several activities may correspond to the movements of only certain body joints. This fact also applicable to the FD which is included in the core building blocks of systems under the umbrella of automatic Human Activity Recognition (HAR).

Tower this end, several recent studies [11, 31, 32, 34] used skeleton joints to detect different scenarios of fall. C. Lee and V. Lee [11] use the Kinect v1 camera as the input sensor and Microsoft Kinect SDK to get skeleton data for their FD system. They chose to track the hip center joint whose position and velocity are used to detect three scenarios: fall in open space from

[9] UR Fall Detection Dataset: http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html

walking or standing, fall from lying in bed, and non-fall events. Their system achieved a 90% FD accuracy fall because they could not detect falls off the chair.

Christopher, Li, and Chung [31] developed algorithms to detect fall and also using Microsoft Kinect v1. They used the positions of all the 20 joints offered by Kinect v1 to calculate the floor plane equation and average velocity to detect falls. Using all joints in their system can distinguish between falls and slowly lying down on the floor. However, their system has problems detecting falls in cases like when somebody jumps in front of the camera; when a person walks out of the Kinect's vision area, it is occasionally detected as a fall; etc. Also, this algorithm does not perform very well on stairs.

In 2015, Bian et al. [32] propose a robust FD approach based on human body part tracking using a single depth camera. In their scheme, a pose-invariant Randomized Decision Tree (RDT) algorithm is proposed for the 3D body joint extraction to capture the fall motion. So, they extract the body joints using their own algorithm which applied to the depth frames not the skeleton frames which already offered by the Kinect as our work. And since the head and hip are the most visible body parts, they use body part classification to extract them. To test their method, the SVM classifier employed on fall, and some ADL (crouching down, standing up, sitting down, walking) to determine whether a fall motion occurs.

The most recent related work in fall features was published in 2016 by Maldonado et al.[34]. The authors investigated feature selection to identify the features that distinguish most between fallen and non-fallen poses. They included three poses in their dataset: walking, sitting and fallen. They used the Kinect depth camera as the input sensor and preprocessed the input data using background subtraction to detect the difference cloud of points that represent the person. Afterward, from the cloud points, they calculated twenty features describing falls. Finally, they used two methods to select the most relevant features: Genetic Algorithm (GA) and Principal Component Analysis (PCA). GA produced six features that best identify falls: three features represent angles, two represent the bounding box, and one is a moment invariant.

Overall, the above works used either a fixed set or all of the skeleton joints; however, none of them tried to determine the joints which are most effective in detecting various types of fall scenarios. Maldonado et al. [34] tested some features that can be extracted from depth data (not skeleton data). Researchers must investigate the 25 skeleton joints among those offered by Kinect v2 (SDK 2.0) as in Figure 4.b and their features that most efficiently could detect various fall scenarios.

4. Discussion And Conclusion

In this survey paper, author introduces a brief background about the most popular device- based fall detection approaches in general. Also overviews the acquisition sensor (the Kinect camera), its hardware and software. Then showed the skeleton tracking using Kinect. Also, the reasons that make researchers choose the Kinect camera and its official SDK to detect the fall of the physically-disabled cane users are explained. The related works on FD that used Kinect are overviewed; their limitation\disadvantages and how to solve them. Followed by a discussion of the existing Kinect-based fall datasets. Finally, the skeleton joints and features that used for FD are overviewed.

There are still many issues and challenges that motivate the development of new Kinect-based FD technique to improve the accuracy under more realistic conditions. Some of these issues for society and the health care system that result from increasing disabling population such as increase in health care costs, shortage of caregivers, dependency and larger impact on society [1].

And the rest issues were disadvantages in the literature techniques of FD [11-13, 33]. These techniques have designed to monitor the healthy people or home alone elderly, but not specialty for the people with physical disabilities who use canes. And some of them [11], was lacking to detect some postures such as falls off the chair which is an important posture to be detected for the disabled people who use chair frequently to rest. While other [12] has a high false alarm rate. And some [33] trained their system with the less natural dataset. Yet some [13] have low accuracy for sitting and lying positions, and when the faller is far from the Kinect. But the main issue in FD is how to distinguish between a fall and other daily activities which are very similar to a fall [30].

As for FD datasets recorded using the Kinect camera, our study revealed that many of them [11-13, 29], differ in terms of their coverage, size stream types, and availability. This also highlights a need for a fall dataset that is available to researchers, that covers many fall types and scenarios, and most importantly, that is recorded in all the streams provided by the Kinect camera to provide for the needs of different techniques.

Furthermore, for the skeleton-based fall detection related works [11, 30, 31, 34], they used either a fixed set or all of the skeleton joints; however, none of them tried to investigate the joints which are most effective in detecting various types of fall scenarios.

The research on FD can be extended from dataset aspects and the proposed method aspects, and the implementation aspects:

1. From the dataset aspects, the existing datasets could be improved by feeding more scenarios in dark rooms or late evenings. And with more female and male subjects who have different body shapes and sizes.
2. From the proposed method aspects, they can be improved by solving the false positive scenarios, provide it by a postural recognition algorithm as done in [11] which could be applied after the detection to test if they are real fall or just a false positive.
3. From the implementation aspects, the proposed methods must develop a completed FD system; and feed it with voice recognition algorithm in order to validate the fall and reduce the false positive as done in [31].

REFERENCES

[1] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," Biomedical and Health Informatics, IEEE Journal of, vol. 17, pp. 579-590, 2013.

[2] E. Melis, R. Torres-Moreno, H. Barbeau, and E. Lemaire, "Analysis of assisted-gait characteristics in persons with incomplete spinal cord injury," Spinal Cord, vol. 37, pp. 430-439, 1999.

[3] E. Dean and J. Ross, "Relationships among cane fitting, function, and falls," Physical therapy, vol. 73, pp. 494-494, 1993.

[4] L. P. Fried, L. Ferrucci, J. Darer, J. D. Williamson, and G. Anderson, "Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care," The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, vol. 59, pp. M255-M263, 2004.

[5] M. Mubashir, L. Shao, and L. Seed, "A survey on fall detection: Principles and approaches," Neurocomputing, vol. 100, pp. 144-152, 2013.

[6] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," ACM Computing Surveys (CSUR), vol. 46, p. 33, 2014.

[7] R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," Biomed. Eng. Online, vol. 12, pp. 1-66, 2013.

[8] Z. Zhang, C. Conly, and V. Athitsos, "A survey on vision-based fall detection," in Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments, 2015, p. 46.

[9] N. Noury, A. Fleury, P. Rumeau, A. Bourke, G. Laighin, V. Rialle, et al., "Fall detection-principles and methods," in 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pp. 1663-1666.

[10] Y. Li, D. Zhou, X. Wei, Q. Zhang, and X. Yang, "Key Frames Extraction of Human Motion Capture Data Based on Cosine Similarity," presented at the The 30th International Conference on Computer Animation and Social Agents (CASA 2017), Seoul, South Korea, 2017.

[11] C. K. Lee and V. Y. Lee, "Fall detection system based on kinect sensor using novel detection and posture recognition algorithm," in Inclusive Society: Health and Wellbeing in the Community, and Care at Home, ed: Springer, 2013, pp. 238-244.

[12] B. Kwolek and M. Kepski, "Human fall detection on embedded platform using depth maps and wireless accelerometer," Computer methods and programs in biomedicine, vol. 117, pp. 489-501, 2014.

[13] E. E. Stone and M. Skubic, "Fall detection in homes of older adults using the Microsoft Kinect," Biomedical and Health Informatics, IEEE Journal of, vol. 19, pp. 290-301, 2015.

[14] F. Abdali-Mohammadi, M. Rashidpour, and A. Fathi, "Fall Detection Using Adaptive Neuro-Fuzzy Inference System," International Journal of Multimedia and Ubiquitous Engineering, vol. 11, pp. 91-106, 2016.

[15] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," Communications Surveys & Tutorials, IEEE, vol. 15, pp. 1192-1209, 2013.

[16] M.-C. Giuroiu and T. Marita, "Gesture recognition toolkit using a Kinect sensor," in Intelligent Computer Communication and Processing (ICCP), 2015 IEEE International Conference on, 2015, pp. 317-324.

[17] Z. Zhang, "Microsoft kinect sensor and its effect," MultiMedia, IEEE, vol. 19, pp. 4-10, 2012.

[18] H.-H. Wu and A. Bainbridge-Smith, "Advantages of using a Kinect Camera in various applications," University of Canterbury, 2011.

[19] Z. Cai, J. Han, L. Liu, and L. Shao, "RGB-D datasets using microsoft kinect or similar sensors: a survey," Multimedia Tools and Applications, pp. 1-43, 2016.

[20] B. Langmann, K. Hartmann, and O. Loffeld, "Depth Camera Technology Comparison and Performance Evaluation," in ICPRAM (2), 2012, pp. 438-444.

[21] M. L. Anjum, O. Ahmad, S. Rosa, J. Yin, and B. Bona, "Skeleton tracking based complex human activity recognition using kinect camera," in Social Robotics, ed: Springer, 2014, pp. 23-33.

[22] J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced computer vision with microsoft kinect sensor: A review," Cybernetics, IEEE Transactions on, vol. 43, pp. 1318-1334, 2013.

[23] J. Smisek, M. Jancosek, and T. Pajdla, "3D with Kinect," Consumer Depth Cameras for Computer Vision, pp. 3-25, 2013.

[24] S. Zennaro, M. Munaro, S. Milani, P. Zanuttigh, A. Bernardi, S. Ghidoni, et al., "Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications," in Multimedia and Expo (ICME), 2015 IEEE International Conference on, 2015, pp. 1-6.

[25] C. Sinthanayothin, N. Wongwaen, and W. Bholsithi, "Skeleton Tracking using Kinect Sensor & Displaying in 3D Virtual Scene," International Journal of Advancements in Computing Technology, vol. 4, 2012.

[26] K. K. Biswas and S. K. Basu, "Gesture recognition using microsoft kinect®," in Automation, Robotics and Applications (ICARA), 2011 5th International Conference on, 2011, pp. 100-103.

[27] Z. Sun, S. Tang, H. Huang, Z. Zhu, H. Guo, Y.-e. Sun, et al., "SOS: Real-time and accurate physical assault detection using smartphone," Peer-to-Peer Networking and Applications, 2016.

[28] Z. Zhu, H. Guo, and Y.-e. Sun, "iProtect: Detecting Physical Assault Using Smartphone," in Wireless Algorithms, Systems, and Applications: 10th International Conference, WASA 2015, Qufu, China, August 10-12, 2015, Proceedings, 2015, p. 477.

[29] G. Mastorakis and D. Makris, "Fall detection system using Kinect's infrared sensor," Journal of Real-Time Image Processing, vol. 9, pp. 635-646, 2014.

[30] F. Abdali-Mohammadi, M. Rashidpour, and A. Fathi, "Fall Detection Using Adaptive Neuro-Fuzzy Inference System," 2016.

[31] C. Kawatsu, J. Li, and C. Chung, "Development of a fall detection system with Microsoft Kinect," in Robot Intelligence Technology and Applications 2012, ed: Springer, 2013, pp. 623-630.

[32] Z.-P. Bian, J. Hou, L.-P. Chau, and N. Magnenat-Thalmann, "Fall detection based on body part tracking using a depth camera," IEEE journal of biomedical and health informatics, vol. 19, pp. 430-439, 2015.

[33] T.-L. Le and J. Morel, "An analysis on human fall detection using skeleton from Microsoft Kinect," in Communications and Electronics (ICCE), 2014 IEEE Fifth International Conference on, 2014, pp. 484-489.

[34] C. Maldonado, H. Ríos, E. Mezura-Montes, and A. Marin, "Feature selection to detect fallen pose using depth images," in Electronics, Communications and Computers (CONIELECOMP), 2016 International Conference on, 2016, pp. 94-100.

Author

Salma Kammoun Jarraya received a Ph.D in Computer Science from Sfax University, Tunisia. She is a researcher in the MIRACL laboratory (Multimedia, InfoRmation systems and Advanced Computing Laboratory). Currently, she is Assistant Professor in computer science, CS Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, KSA. Her research interests include computer vision, video and image processing. She has served on technical conference committees and as reviewer in many international conferences and journals.