

Article

Strategies to Improve Activity Recognition Based on Skeletal Tracking: Applying Restrictions Regarding Body Parts and Similarity Boundaries [†]

Carlos Gutiérrez-López-Franca *, Ramón Hervás and Esperanza Johnson

MAMI Research Lab, University of Castilla-La Mancha, Paseo de la Universidad 4, 13071 Ciudad Real, Spain; ramon.hlucas@uclm.es (R.H.); mesperanza.johnson@alu.uclm.es (E.J.)

* Correspondence: carlos.gutierrez.uclm@gmail.com

[†] This paper is an extended version of our paper published in Gutiérrez López de la Franca, C.; Hervás, R.; Johnson, E.; Bravo, J. Findings about Selecting Body Parts to Analyze Human Activities through Skeletal Tracking Joint Oriented Devices. In Proceedings of the 10th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAMI 2016), Gran Canaria, Spain, 29 November–2 December 2016.

Received: 4 April 2018; Accepted: 17 May 2018; Published: 22 May 2018



Abstract: This paper aims to improve activity recognition systems based on skeletal tracking through the study of two different strategies (and its combination): (a) specialized body parts analysis and (b) stricter restrictions for the most easily detectable activities. The study was performed using the *Extended Body-Angles Algorithm*, which is able to analyze activities using only a single key sample. This system allows to select, for each considered activity, which are its *relevant joints*, which makes it possible to monitor the body of the user selecting only a subset of the same. But this feature of the system has both advantages and disadvantages. As a consequence, in the past we had some difficulties with the recognition of activities that only have a *small* subset of the joints of the body as relevant. The goal of this work, therefore, is to analyze the effect produced by the application of several strategies on the results of an activity recognition system based on *skeletal tracking joint oriented devices*. Strategies that we applied with the purpose of improve the recognition rates of the activities with a small subset of relevant joints. Through the results of this work, we aim to give the scientific community some first indications about which considered strategy is better.

Keywords: activity recognition; Kinect; ubiquitous computing; ambient intelligence; extended body-angles algorithm

1. Introduction

The analysis and recognition of human activities is a very popular area of work and research. Due to this fact, there are an innumerable amount of proposals in this research area with several types of focus and devices: machine learning methods, video analysis, mobile devices, wearable sensors and many more. Even though there are a variety of approaches when attempting to overcome the problem of activity recognition, there are some similarities as well. For example, most of the existing works are usually based on machine learning techniques.

On the other hand, it is not as typical to focus on analyzing certain aspects or strategies applicable to different systems, with the goal of improving recognition rates. As an example, there is the analysis of the effect that different body parts have on the results of activity recognition systems [1]. Though some of the existing works focus their analysis on *specific body parts* [2], it is not common to evaluate the existing differences between using some parts of the body or not [3,4]. These precedents demonstrate that it is interesting to center the analysis *only on specific parts of the body*.

The *Extended Body-Angles Algorithm (E-BA-A)* [5] is an activity recognition system that allows one to compare the similarity between two *postures* or *movements* performed by human bodies. This system, following the taxonomy of David and Marcus [6], is considered a case-based commonsense reasoning approach based on mathematical analogies. This kind of algorithm has only very limited interaction with machine learning approaches, that typically work over large data corpora. In fact, the main feature that differentiates this activity recognition system from others is that the *E-BA-A* doesn't need to use a dataset with multiple instances of each supported activity. It is capable of performing the analysis with a *single instance* thanks to the use of the *angles* that are formed by each pair of joints as its unit of information. Another of its main features is the possibility to personalize which body parts are relevant in each activity. This way, we can optimize the computational cost by frame of the algorithm, omitting the processing of the irrelevant body parts. Taking into account the complexity of the *E-BA-A*, its functioning is explained more in depth in a previous publication [5]. In it, we also cover the aspects related to the comparison between the *E-BA-A* and the most common used recognition methods (based on machine learning techniques).

The goal of this work is to analyze the effect that is produced by the application of several strategies on the results of an activity recognition system. Specifically, on an activity recognition system based on the use of *skeletal tracking joint oriented devices*. As its name suggests, these devices perform the skeletal analysis tracking the 3D-position of the joints of the body. Therefore, this kind of device outputs the position in the environment of each one of the supported joints. We have used *Microsoft Kinect* as our sensor device and the *E-BA-A* as the activity recognition system in which we have applied the considered strategies of improvement. The extrapolation to other systems will be addressed in the discussion (Section 7).

In order to achieve this goal, in first place, we will expose which were the circumstances that led us to need to apply an improvement strategy in this kind of system. Later we will detail each one of the strategies applied, as well as the obtained results in each evaluation. The improvement strategies we have considered in this study are:

- (1) Study how the use of different body parts affects the results in order to select as relevant only those joints that contribute to the improvement of the recognition.
- (2) Apply greater restrictions (establish a minimum similarity limit that has to be surpassed) to those activities that are detected most easily through skeletal tracking joint oriented devices.
- (3) Combine the use of the two previous strategies.

The article has following structure: Section 2 will explore in depth the contemplated state of the art; Section 3 will expose the difficulties that motivated the search of strategies to improve recognition rates; Sections 4–6 will describe each one of the strategies applied and the tests that have been performed (usually following the structure: *Description of the Strategy*, *Description of the Strategy Evaluation*, *Evaluation Results* and *Analysis of the Results*); the discussion will be presented in Section 7; and finally Section 8 will cover the final conclusions extracted from this study and the future work. In addition, the article has the following appendixes: Appendix A considers activities during the performance of the evaluations; Appendix B gives the body information of the volunteers; Appendix C describes the hardware used during the evaluations; and Appendix D covers particularities of the evaluations.

2. State of the Art

There are a wide range of proposals centered on the area of *activity recognition*, as it is a topic of study that generates a lot of interest. There are proposals to evaluate the performance of a dancer [7], to allow physical rehabilitation [8], activity recognition for the physical rehabilitation through mobile devices with accelerometer [9], detection of elderly frailty with accelerometry [10], detection of stereotypical behavior in autistic children [11], a review of human activity recognition methods and different examples of each method [12], an analysis of the current outlook of research in

the field of activity recognition [13] and a system that monitors the activities that occur in an isolation room of a psychiatric hospital [14].

There are also proposals that support themselves on *objects from the environment* (taking context into account) to do the activity analysis. Some of those examples are monitoring the activities and state of a dog through sensors embedded in their collar [15], a system of activity recognition of daily living that takes context into account (the objects the user interacts with) [16] and lastly the detection of activities done in a house with simple ubiquitous sensors [17].

Other kinds of proposals center around the use of *alternative recognition strategies*, like the recognition of activities in which the observer in 1st person takes part [18], a study of activity recognition when information from skeleton tracking is not available [19], or gesture recognition [20] and hand gesture [21].

Some of the studied works analyze activities focusing only on a *specific body part*. Some examples are a system to guide people with dementia in the performance of activities of daily living [22], a system based on context centered around culinary activities and objects [23] and a method capable of recognize activities from a video camera in an egocentric point of view analyzing the hands of the user and the objects [2]. But, in this group of works, those that approach the possible differences in recognition are mostly related with *variations in the disposition of the accelerometers* on the human body. There are studies that analyze the impact that these variations of the accelerometers from their initial position in which they were put have [24,25], studies about the *optimal placement position of accelerometers* and how to improve recognition combining the data from several of them [3] and a method capable of *obtaining the position of the body* in which an accelerometer has been placed based only on the signal of the device [26] (as knowing where the sensor is located is very important information about context).

Lastly, most similar related work to our proposal corresponds to two proposals opposite to each other. The first of them proposes to analyze human activities distinguishing among the different body parts of the person and applying a relational representation by means of graphs [4]. This way they attempt to test if this representation obtains better results than considering the whole body as a single object. The second proposal aims for the opposite [27] and maintains that it is possible to recognize repetitive activities without discerning between the different parts that compose the human bodies. For this, they present a method capable of performing this analysis through sequences of images in greyscale.

Regarding the application of greater restrictions on the most common activities and the ones most easily detectable, we are not aware of any related work. Therefore, during the study of the state of the art, we have not found any precedent whose focus is exactly the same as this proposal: compare the effect of different improvement strategies on activity recognition systems. In consequence, a comparative analysis can be beneficial for future or existing works whose recognition is based on *skeletal tracking joint oriented devices*.

3. Difficulties of the Activity Analysis Based on Relevant Joints

In the evaluations previously done to the *E-BA-A* [5] we identified different types of activities that, due to their nature, are more difficult to monitorize. With the skeletal tracking joint oriented devices, depending of what kind of joint selection is made, the activity is more or less susceptible to monitoring errors. From a general perspective we can identify two traits to classify activities:

- *Global/Bounded Use of the Body*: The activity is characterized by using all or most of the monitored joints (*Global*) or by using a reduced subset of the same (*Bounded*).
- *Symmetry*: There are activities that use both halves of the body, either symmetrically or with an inverted symmetry, while others are characterized by using only one half of the body.

According to the monitored activities and the results of the evaluation, the movements that use the joints in a *global* way (like *Walk* or *Forward Bend*) obtained very good accuracy percentages. But in the case of activities that only took into account one subset (*bounded*), like for example the joints of

the right arm, the results were significantly worse. This is mainly due to two factors: the number of relevant body parts and the postural coincidences.

The first factor is the influence that considering a higher or lower number of joints as relevant has on the *E-BA-A*. If an activity contemplates a greater number of joints, the effect that each joint can have individually, over the total result, is lesser. Therefore, the activities with less relevant joints are more exposed to errors. We will analyze this factor in depth during the development of Section 4, dedicated to the first of the considered strategies.

The second factor is the *postural coincidences* between different activities. There are different activities in which the position of certain joints and bones of the body is the same. When these similarities happen it makes it more difficult to distinguish the activities. This situation becomes specially critical when these *postural coincidences* occur between an activity in which the coincidental part of the body is relevant, and an activity in which it is not relevant.

These problems have motivated the search for strategies that would allow us to improve the results of the *E-BA-A*. During the next sections we will describe in detail the strategies that we have applied and analyzed. We will present how they have affected the results obtained by the *E-BA-A* in the evaluations that we have performed.

4. Strategy #1—Specialized Body Parts Analysis

4.1. Description of the Strategy

The possibility of *defining a personalized set of relevant joints for each activity*, gives the *E-BA-A* both positive and negative effects. By this reason, this feature should be used with precaution. We can summarize the effects in:

- *Positive Effects*
 - *Efficiency of the analysis*—Consider as relevant only a subset of the joints, reduces the computational load that the *E-BA-A* needs to analyze that activity. A smaller number of relevant bones leads to fewer mathematical calculations.
 - *Flexible use of the samples*—Using only the subset of relevant joints for the analysis, the *E-BA-A* can monitor the activity with a single sample of it, independently of the position of the rest of the joints. But if we consider the irrelevant joints too, it is possible that we will need several samples of the same activity depending on the positioning possibilities of the irrelevant joints.
- *Negative Effects*
 - *Each joint has greater influence*—If there are less *relevant bones*, each bone similarity has a greater impact in the final result. This fact could be positive or negative depending on the similarity value obtained. With less relevant bones, high bone similarity values benefits more the determination of high *body similarities*, but low bone similarities penalize much more the final result. This implies that any activity that has less relevant bones will be more exposed to possible precision errors that the sensor device can make.
 - *Avoid stable parts of the body*—There are certain parts of the body whose tracking is easier because its movement possibilities are limited (mainly the *trunk*). An easier tracking means an extra of *stability* (less vulnerable to precision errors) in the results. An activity with a reduced subset of relevant joints that do not consider those stable body parts, will be at a certain disadvantage.

Taking into account the negative effects, we could consider eliminating this feature (the possibility of *defining a personalized set of relevant joints for each activity*) as a possible solution to the problem (Section 3). But the inconvenience of this option is that we lose the interesting advantages that the

current implementation of the *E-BA-A* offers, without discerning what effects (positive ones or negative ones) have more impact on the system. For this reason, we chose an intermediate approach for the first strategy applied.

The first strategy proposed, to improve the detection of the activities with a bounded use of the body, consists in: *studying the influence that body parts have on the recognition results*. The goal behind this intermediate approach is trying to reduce the negative effects that this characteristic of the algorithm introduces, but without rejecting the positive ones. If we can identify parts of the body that maximizes the obtention of high similarities, we can assign the *joints* of these parts as *relevant* (for the *bounded* activities).

To perform this study, we have defined several groups of different joints. As a matter of fact, one of the groups that has been contemplated in the evaluation is the complete set of supported joints (vC). We are attempting to reach a less radical solution than to use all (supported) joints of the body for all the activities. But we have chosen to include this scenario as well in order to study if its results (in comparison with the other groups of joints) clearly support its application.

In the next subsection we will explain what each group of joints consists of and what its goal is, as well as the evaluation dynamic of this improvement strategy.

4.2. Description of the Strategy Evaluation

The goal after this test is to find out if, on the activity recognition systems oriented towards *monitorization of joints*, it is more convenient to consider all or only those joints important for each activity. To study if it is more beneficial to include or omit the irrelevant joints, we will compare the recognition results obtained taking into account three different sets of joints.

The first set (vB) only considers the *joints of the extremities* used in the movement *without including the hands* (when the movement is done with the arm/s) *and the feet* (when the movement is done with the foot/feet). We omit these elements because in previous work, working with Kinect too, we observed that they tend to worsen the results. This way we want to check if just the elimination of these elements is enough to improve the recognition results of the movements whose relevant joints are a reduced subset. The joints corresponding to the shoulders and hips are also part of the set vB, as the union points between the trunk and the extremities belong to both.

The second set (vT) considers the same joints as the set vB but adding those belonging to the *trunk*. With this we try to test if considering the trunk (as it is one of the parts with a more stable tracking) is viable to obtain high rates of accuracy without giving up the possibility of using samples of these movements from different positions. It is a compromise engagement, an option that involves more joints but is not as radical as to stop omitting irrelevant joints.

The last set (vC) corresponds to *all the joints of the body* except hands, feet and head. Use of all the body theoretically eases the differentiation between activities. Therefore, with this set we try to test if accuracy rates in recognition improves remarkably in comparison with sets vB and vT. Figure 1 shows a visual representation of the three different sets, with the relevant bones in each case coloured in yellow.

As we have considered three different sets of joints, we have three different versions for each considered activity (each one of them with the relevant joints defined by each set). The four selected activities are: *Drink [D]* and *Grab Object [D]*, from the previous evaluation (and done with the arms), and *Leg Flexion [I]* and *Leg Circular Swing [I]* as new additions (that use the legs as the extremity). The activities identified with [D] are done with the right side of the body. These are the *asymmetric activities*, those that can be performed with either the left [I] or right side of the body, and therefore two versions must be recorded in the system, one for each half of the body.

One aspect that would be convenient to clarify is the *reason to select these activities*. Since the aim of this work is not to apply the activity recognition to a particular scenario, for a better understanding is important to remark that this selection is motivated for previous works about analyzing IADL activities in elderly people with and without dementia. One of the typical ways to evaluate cognitive

impairment is through a *dual task*. The elder person performs one of those physical activities while doing some cognitive activity (for example, counting). But we selected these activities because we are looking for the classification improvement of those activities more difficult to detect through the skeletal tracking technology. Taking into account the results of previous work [5] these are the bounded activities (only a small subset of the joints is relevant). These four selected activities are good examples of activities difficult to classify for the *E-BA-A*. And we choose these four to have two of them performed with the arms and two of them performed with the legs. Besides, its relevant joints correspond only to the *extremities*. In consequence, they are ideal for the goals of this evaluation as they are compatible with the set *vB*.

The rest of the details of the evaluation can be seen in Appendixes A–D, as these are aspects common to the rest of the evaluations performed in this work. Some of the aspects detailed in the appendixes are: the information about the bodies of the volunteers, the hardware used during the evaluations or the rest of the activities considered by the *E-BA-A* (apart from the selected ones for the evaluation).

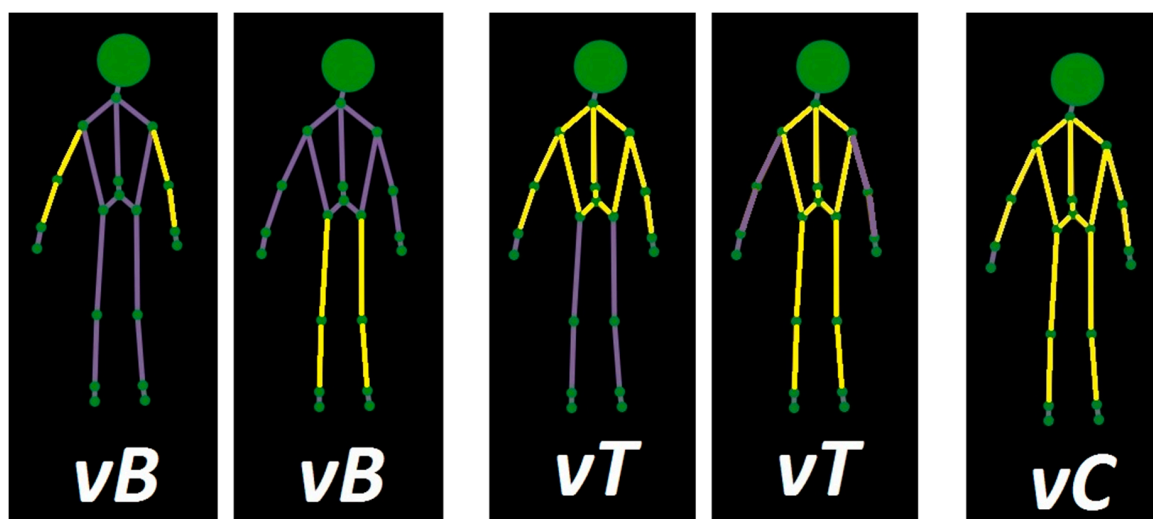


Figure 1. Visual representation of the sets *vB*, *vT* and *vC*.

4.3. Evaluation Results

Figure 2 shows the results obtained in this first evaluation. We can distinguish different types of information including the results from each volunteer for each of the movements, in addition to the totals by movement and totals per volunteer.

Each unit of information contains four fields. In first place, *Suc* indicates the number of times that the algorithm predicted correctly the movement performed by the participant. In second place, *Vic* indicates the number of times that that set of joints (*vB*, *vT* or *vC*) has obtained a similarity with the activity greater than the rest of considered sets. Both fields are accompanied by its corresponding success rate (*Rate*). In this first evaluation each test subject accomplished all activities a total of 50 times.

Specialized Body Parts Analysis Evaluation – Strategy #1 – Test Results – Sensors 2017																				
vB	Drink [D]				Grab Object [D]				Leg Flexion [I]				Leg Circular Swing [I]				Total			
	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate
Subject #1	49	98 %	0	0 %	39	78 %	4	8 %	0	0 %	0	0 %	10	20 %	3	6 %	98	49 %	7	3.5 %
Subject #2	45	90 %	0	0 %	50	100 %	1	2 %	0	0 %	0	0 %					95	63.33 %	1	0.667 %
Subject #3	50	100 %	0	0 %	50	100 %	48	96 %	0	0 %	0	0 %	0	0 %	0	0 %	100	50 %	48	24 %
Subject #4	50	100 %	0	0 %	50	100 %	10	20 %	0	0 %	0	0 %	8	16 %	0	0 %	108	54 %	10	5 %
Subject #5	47	94 %	0	0 %	50	100 %	44	88 %	0	0 %	0	0 %	0	0 %	0	0 %	97	48.5 %	44	22 %
Total	241	96.4 %	0	0 %	239	95.6 %	107	42.8 %	0	0 %	0	0 %	18	9 %	3	1.5 %	498	52.42 %	110	11.58 %
vT	Drink [D]				Grab Object [D]				Leg Flexion [I]				Leg Circular Swing [I]				Total			
	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate
Subject #1	50	100 %	49	98 %	49	98 %	45	90 %	50	100 %	49	98 %	40	80 %	40	80 %	189	94.5 %	183	91.5 %
Subject #2	50	100 %	50	100 %	50	100 %	49	98 %	49	98 %	15	30 %					149	99.33 %	114	76 %
Subject #3	50	100 %	49	98 %	50	100 %	2	4 %	45	90 %	15	30 %	50	100 %	41	82 %	195	97.5 %	107	53.5 %
Subject #4	50	100 %	50	100 %	50	100 %	40	80 %	48	96 %	10	20 %	49	98 %	10	20 %	197	98.5 %	110	55 %
Subject #5	50	100 %	50	100 %	50	100 %	6	12 %	49	98 %	0	0 %	46	92 %	12	24 %	195	97.5 %	68	34 %
Total	250	100 %	248	99.2 %	249	99.6 %	142	56.8 %	241	96.4 %	89	35.6 %	185	92.5 %	103	51.5 %	925	97.37 %	582	61.26 %
vC	Drink [D]				Grab Object [D]				Leg Flexion [I]				Leg Circular Swing [I]				Total			
	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate
Subject #1	50	100 %	1	2 %	36	72 %	0	0 %	50	100 %	1	2 %	29	58 %	0	0 %	165	82.5 %	2	1 %
Subject #2	50	100 %	0	0 %	50	100 %	0	0 %	49	98 %	35	70 %					149	99.33 %	35	23.33 %
Subject #3	50	100 %	1	2 %	50	100 %	0	0 %	48	96 %	34	68 %	21	42 %	9	18 %	169	84.5 %	44	22 %
Subject #4	50	100 %	0	0 %	50	100 %	0	0 %	50	100 %	40	80 %	50	100 %	40	80 %	200	100 %	80	40 %
Subject #5	50	100 %	0	0 %	47	94 %	0	0 %	50	100 %	50	100 %	48	96 %	36	72 %	195	97.5 %	86	43 %
Total	250	100 %	2	0.8 %	233	93.2 %	0	0 %	247	98.8 %	160	64 %	148	74 %	85	42.5 %	878	92.42 %	247	26 %

Figure 2. Specialized Body Parts Evaluation results for the sets vB-vT-vC.

4.4. Analysis of the Results

The conclusions of this first evaluation were clear. The best approach in order to obtain better recognition results was the use of the set vT, meaning including the *trunk* as a relevant part of the body in all the activities. The set vC obtained high accuracy rates too (92.42%) but not as good as the results of the set vT (97.37%). Finally, the accuracy rates obtained by the set vB were far from the results obtained by vC and vT. The set vB obtained a global success rate of 52.42% highlighting the high contrast between the results obtained in the recognition of the activities performed with the arms (96.4% for *Drink [D]* and 95.06% for *Grab Object [D]*) and the activities performed with the legs (0% and 9% respectively for *Leg Flexion [I]* and *Leg Circular Swing [I]*).

5. Strategy #2—Stricter Restrictions for the Most Easily Detectable Activities

5.1. Description of the Strategy

The fact that the algorithm fails, confusing the performed activity with another of the most easily detectable, was already planned that it may occur [5]. The most basic and common activities, such as *Stand*, will have many joints positions in common (*postural coincidences*) with others. These kind of activities is what we call *sink* postures or movements. We selected the term *sink* to identify these activities due to the effect they produce on the results. They absorb samples as their own when they should be classified as a different activity. In other words, the *sink* activities obtain high percentages of similarity when the user is actually performing another activity.

These *postural coincidences* (Section 3) are particularly determinant when they occur with *bounded* activities. But *postural coincidences* are not the only factor that make it easier to obtain better results with *global* activities instead of *bounded* activities. The factors are listed below:

- *Postural coincidences*
 - *Irrelevant joints for an activity could be relevant for another*—Even though the information of the irrelevant joints of an activity is ignored by the *E-BA-A* in its analysis, these joints

will be relevant for other activities. Therefore, it is possible that the position of these joints (irrelevant in one case and relevant in other) will be the same in both cases. This situation usually happens, thus the detection and differentiation between activities is more difficult. This is especially critical with the *bounded* activities because they have less joints to differentiate itself from the rest of the activities.

- *Example*—To illustrate this we can describe as an example the activity *Grab Object*. This is an activity bounded with a high significance to the joints of the arm. If this activity is performed standing up, it can be easily confused with the activity *Stand* as the position for the rest of the joints is almost the same.
- *Greater exposure to precision errors* (Section 4.1 “Each joint has greater influence”)
- *Limitations of the sensor device*
 - *Computer vision devices nature*—The skeletal tracking joint oriented devices, due to their nature, are not capable of performing the tracking of the user’s body with the same level of precision for all kinds of actions. This means that, depending on the performed activity by the user, the device will have more or less difficulties to capture all the variations that have happened in the position of the joints of their body.
 - *Microsoft Kinect specific case*—Kinect is a device that is comfortable analyzing human bodies that are in front of it and looking to it. This is reasonable, given that the original application with which Kinect was made was to be an input device for the console *Microsoft XBOX 360*. A peripheral that allows the user to become a peripheral themselves. In this context, the interactions made by the users are always explicit because the user must stand in front of Kinect in order to be able to receive the feedback of videogames. But even in this situation Kinect has some limitations: *positioning joints imprecisions* even when the user is still, greater difficulties to *capture sudden variations* of position and the *inference of joints* when they are covered by another body part (like any computer vision device) and Kinect does not “see” them. This limitation has made it so that even some videogames developed by important companies in the sector, exclusive for Kinect, do not have a great reputation. The reason is that those videogames have not achieved a good user experience, because they fail to map the user’s action a significant number of times.

Once we have exposed the three factors that hinder the recognition of *bounded* activities, it is easy to understand the three opposite factors that characterize *sink* activities. These factors are: (1) they use the body in a *global* way; (2) its monitoring is not demanding for the sensor device; (3) if there are *postural coincidences*, these are resolved in their favor, as there is a great chance that those joints will be relevant for the *sink* activity. Of the total number of activities considered (Appendix A) we consider as *sink*: *Stand*, *Stay Seated*, *Walk* and *Walk Backwards*. All of them have in common that they are activities performed smoothly, with a big set of relevant joints and without body parts trajectories that cover others.

Taking all of this into account, we have a group of activities (*sink*) with a greater probability that its elements will be selected as a prediction, and another group (*bounded* activities) with difficulties for being detected. In consequence, the second strategy proposed to improve the detection of the last ones, consists in *applying some sort of restriction to the sink activities*. The goal of this strategy is to reduce the number of occasions in which a *sink* activity is chosen as a prediction by the *E-BA-A* when it is really not being performed. The idea for it is to introduce a barrier. If the barrier value is not surpassed, the system will consider that the *sink* activity is not similar enough to be considered the prediction for that instant (even if it has obtained the higher similarity percentage).

In this respect we have introduced the concept of *minimum limit of prediction* in the past [5]. Its original objective was to discern those situations in which the user was doing an *unknown activity* for the system. So, why not apply this idea on another level? Instead of defining a single limit that has

to be surpassed by any activity to be considered for prediction, establish a stricter and more exclusive limit for sink activities. But, which *value* do we use? *One* or *more* limits? To answer these questions, we have elaborated a preliminary evaluation.

5.2. Description of the Pre-Evaluation

To obtain the necessary data to select the limit to apply, the pre-evaluation is divided in two parts. The first part consists of performing the four sink activities considered (*Stand*, *Stay Seated*, *Walk [D]* and *Walk Backwards [D]*) several times in order to analyze which similarity percentages are obtained approximately when the user is really doing a sink activity. The second part consists of performing the four activities whose recognition rates we are trying to improve (*Drink [D]*, *Grab Object [D]*, *Leg Flexion [I]* and *Leg Circular Swing [I]*) but analyzing which similarity is obtained in each case with each one of the sink activities. This way we can make an approximation of what similarity percentages are obtained by the four sink activities when they are not being performed.

Through the results of both parts of the pre-evaluation we attempt to discover the interval of values in which we can choose the limit or limits to apply. The rest of the details of the pre-evaluation coincide with the rest of the performed tests and can be seen in Appendixes A–D.

5.3. Pre-Evaluation Results

Through this pre-evaluation, what we are looking for is to delimit the intervals of values in which we can choose a limit that could correctly perform our goals. With the performance of the sink activities (second part of the test) we search the upper limit of the interval. The greatest values of similarity that can be reached doing the sink activities. Meanwhile, the similarity values obtained by the sink activities, when the performed activity was not one of them (first part of the test), indicates us the lower limit of the interval. The similarity values that should not surpass the limit if we want to obtain the desired behavior. To ease the comprehension of the limit selection process, Figure 3 shows the results of the pre-evaluation visually.

In our search of the *minimum limit or limits of prediction* we have considered the maximum, minimum and mean of the values that defined the selection interval. The reason for this is that, depending on the size of the interval (the difference between the values that suggest the upper limit and those that suggest the lower limit), we will have to use one of the different statistical values as the definite interval delimiters.

The ideal situation would be that there would be a big interval of values (the wider the better) between the minimum of the upper limit candidates and the maximum of the lower limit candidates. If this happens, we will have more guarantees about the correct functioning of the limit. Meaning, the limit would be surpassed when it has to be surpassed (the user is really performing a *sink* activity), and it wouldn't be surpassed when it hasn't been surpassed (the user is not performing a *sink* activity).

An example of this ideal situation has happened with the results of the sink activity *Stay Seated*. As it can be seen in Figure 3, and with more detail in Figure 4, there is a big range of available values for the activity *Stay Seated*. A big range will facilitate that the selected limit differentiates the cases in which the sink activities are being done, and when they are not. This situation also happens with the activities *Walk [D]* and *Walk Backwards [D]* but with the difference that the available values interval is much smaller.

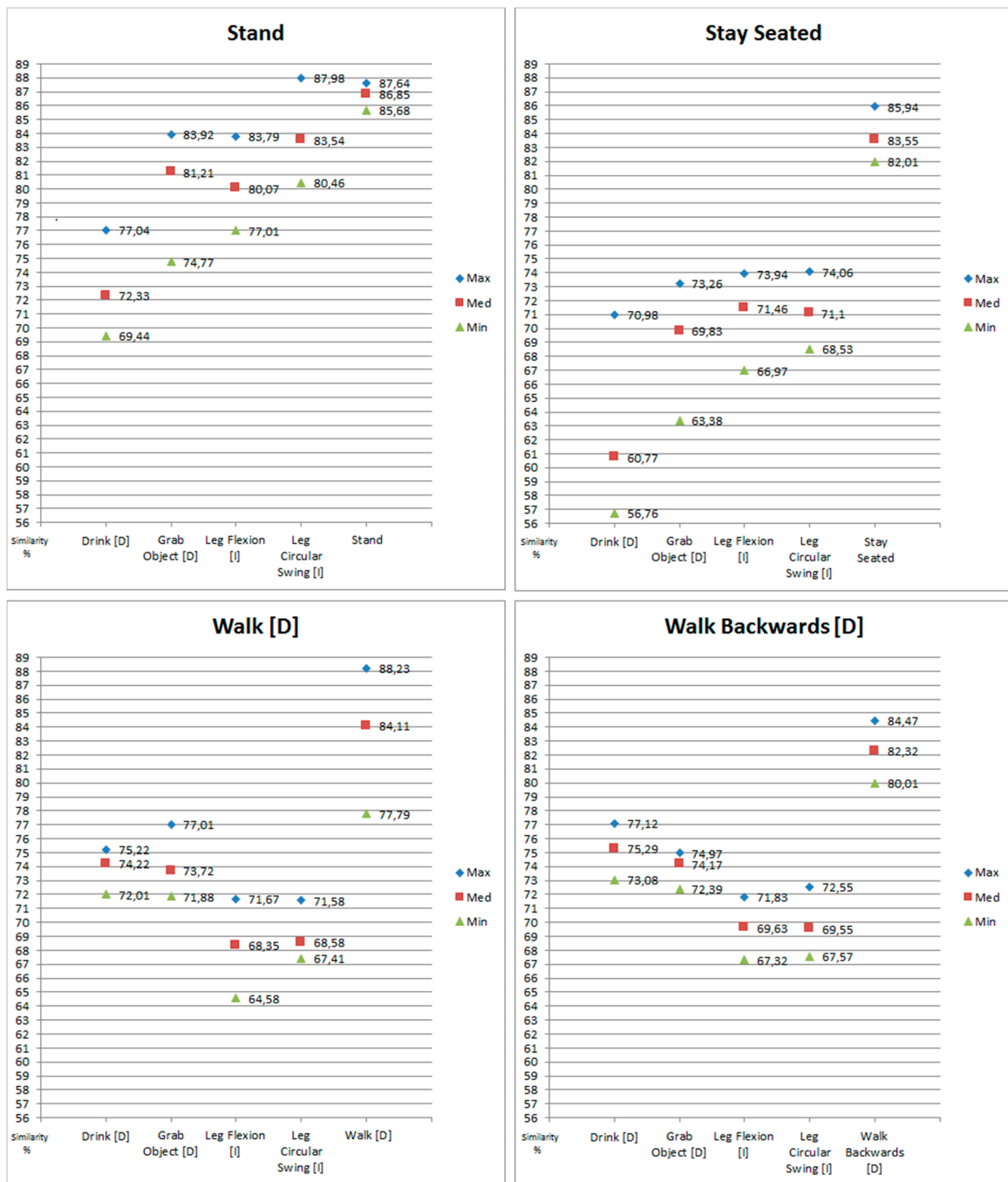


Figure 3. Graphical Representation of the Limit Selection Pre-Evaluation Results.

In the case of *Stand* the situation is more complicated. There is not an interval of values between the minimum of the upper limit candidates and the maximum of the lower limit candidates. This is because the maximum obtained similarity, when the sink activity is not being made, is greater than the minimum obtained similarity when the sink activity was actually being performed. When this happens it is necessary to adjust the selection criteria for the values that will define the selection interval. The consequence is that, even assigning a *minimum limit of prediction*, it will be much more difficult to discern when this activity is really being done and when it is not.

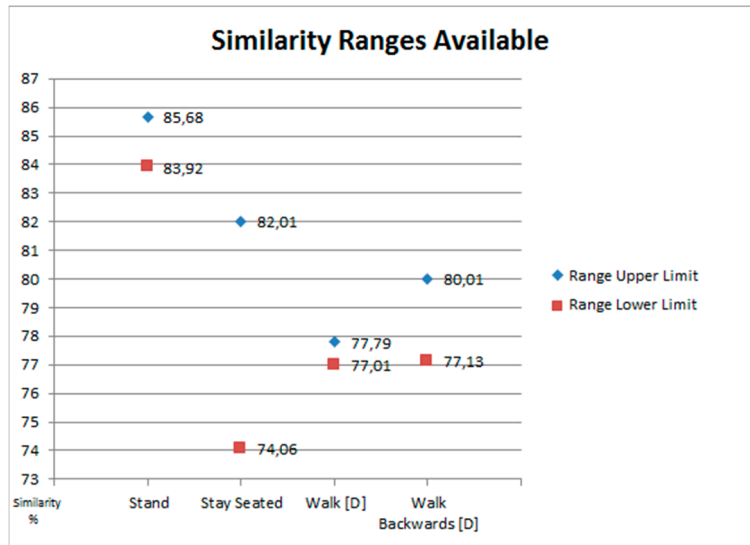


Figure 4. Available ranges of values for selecting *minimum limits of prediction*.

Figure 4 shows the selected upper and lower limits. The *minimum limit of prediction* will be a value that is within the interval defined by those limits. But given the results, it would not be advisable to choose to use a single limit for all four sink activities. In the case of *Stay Seated*, *Walk* and *Walk Backwards* they share a part of the interval’s values. So, it would be feasible to choose a single limit to control all three activities. But in the case of *Stand* its interval is very different from the rest. In consequence, the chosen limit for the other three would not work.

Even if we adjust the selection of the interval for the case of *Stand* using mean values (instead of maximum and minimum to choose the limits), the interval increases but not enough (Figure 5).

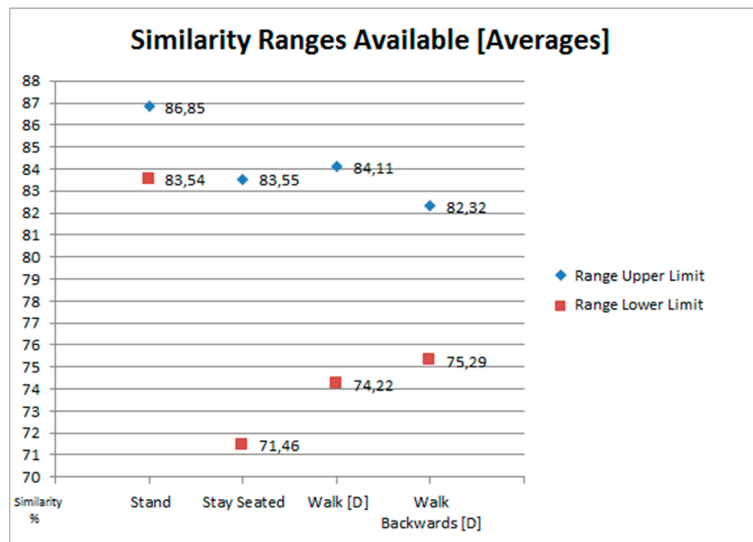


Figure 5. Consequences of widening the range of values of *Stand*.

Taking the situation into account, we decided to use a personalized limit for each sink activity. This way, each limit could work better individually as it is designed only to control a concrete activity. Moreover, with a large number of sink activities, we will probably have the same problem (there are not shared intervals for all cases). Table 1 has the selected values as *minimum limit of prediction* for each one of the *sink* activities analyzed in this pre-evaluation.

Table 1. Minimum Limits of Prediction assigned to the Sink Activities.

Activity	Minimum Limit of Prediction
Stand	85.4%
Stay Seated	78.0%
Walk [D]	78.0%
Walk Backwards [D]	79.0%

5.4. Description of the Strategy Evaluation

Once we have described the preliminary evaluation, the main evaluation of this second strategy is described as follows. With this evaluation we have two objectives:

- (1) *Test if the selected limits are too strict:* Test if even applying a restriction of having to surpass a limit, the sink activities are still being detected when they are performed.
- (2) *Test if the selected limits fulfill their purpose:* Test if the limits prevent that sink activities are selected as a prediction when they are not being performed.

To achieve this, every participant has done a total of 25 instances of both the sink activities and the activities whose recognition we are trying to improve. For this evaluation, we have begun from the base of the results previously obtained by the set vB (Section 4.4). As its results have demonstrated that eliminating hands and feet as relevant joints is beneficial, the samples used in this evaluation do not include these body parts. The rest of the details of the evaluation, common to the rest of the performed tests, can be seen in detail in Appendixes A–D.

5.5. Evaluation Results

Figure 6 shows the results obtained from each participant for each of the activities in addition to the totals by activity and totals per volunteer. Each unit of information (intersection between subject and activity) contains two data. Suc indicates the number of times that the algorithm predicted correctly the movement performed by the voluntary. This field is accompanied by his corresponding success rate (Rate) taking into account that the volunteers performed all activities a total of 25 times (a total of 1000 activities performed during the evaluation).

Stricter Restrictions for the Most Easily Detectable Activities Evaluation – Strategy #2 – Test Results – Sensors 2017													
Activities	Subject #1		Subject #2		Subject #3		Subject #4		Subject #5		Total Results by Activity		Observations
	Suc	Rate	Suc	Rate	Suc	Rate	Suc	Rate	Suc	Rate	Suc	Rate	
Drink [D]	23	92 %	25	100 %	25	100 %	17	68 %	25	100 %	115	92 %	The applied limits have avoided confusions with the sink activities but we found another problem
Grab Object [D]	17	68 %	25	100 %	25	100 %	25	100 %	25	100 %	117	93.6 %	
Leg Flexion [I]	21	84 %	0	0 %	0	0 %	0	0 %	7	28 %	28	22.4 %	
Leg Circular Swing [I]	25	100 %	0	0 %	0	0 %	4	16 %	2	8 %	31	24.8 %	
Total Enclosed Activities	86	86 %	50	50 %	50	50 %	46	46 %	59	59 %	291	58.2 %	
Stand	25	100 %	25	100 %	25	100 %	25	100 %	25	100 %	125	100 %	The use of minimum limits of prediction with sink activities has not prevented its correct detection
Stay Seated	25	100 %	25	100 %	25	100 %	25	100 %	25	100 %	125	100 %	
Walk [D]	25	100 %	25	100 %	24	100 %	25	100 %	25	100 %	124	99.2 %	
Walk Backwards [D]	24	96 %	25	100 %	25	100 %	24	96 %	25	100 %	123	98.4 %	
Total Sink Activities	99	99 %	100	100 %	99	99 %	99	99 %	100	100 %	497	99.4 %	
Total Results by Subject	185	92.5 %	150	75 %	149	74.5 %	145	72.5 %	159	79.5 %	788	78.8 %	

Figure 6. Stricter Restrictions for the Most Easily Detectable Activities Evaluation Results.

Unlike the other strategies, in this second one, we do not use the term Vic. The reason for that is we cannot use Vic in this kind of strategy. Here we have only applied limits to the sink activities. In this strategy, we do not have different sets to compare. As Vic indicates when one set obtains higher similarity than others, it does not make sense.

5.6. Analysis of the Results and Comparative with the Results of the Previous Strategy

We will analyze the effect that the *minimum limit of prediction* have had over the *sink* activities. As it can be seen in Figure 6, the application of these restriction on these types of activities that are easier to detect, has had no negative impact on the results. As a matter of fact, the four activities have had accuracy rates greater than 98%. Even in the case of *Stand* and *Stay Seated* the rate has reached 100%. Analyzing the results from the point of view of the participants, the results have been equally positive, with all participants obtaining an accuracy rate between 99% and 100% while doing the sink activities.

Regarding the *bounded* activities, there has been improvement in the accuracy rates, but not as much as expected. The activities whose relevant joints are only in the arms (*Drink [D]* and *Grab Object [D]*) have gotten very good results (92% y 93.6% of accuracy rate respectively). If we compare these numbers to the ones obtained by these activities through the set vB (96.4% and 95.6% respectively), the results are slightly worse, but the differences are not so great as to be significant. Especially considering that this evaluation has been more aggressive than the evaluation of the first strategy. Performing less instances of the activities, the detection errors penalize the final accuracy rate so much more.

The activities performed with the legs (*Leg Flexion [I]* and *Leg Circular Swing [I]*) have obtained very low accuracy rates (22.4% and 24.8% respectively) but the result is significantly better than the one through set vB (0% and 9% respectively) though that is not improvement enough. It is worth mention the results obtained by the participant #1. An accuracy rate of 84% in *Leg Flexion* and 100% in *Leg Circular Swing*.

Given the results, one would think that the application of limits has not worked as it should have, but this has not been the case. In order to control the development of the evaluation and collect the results, we developed an application. This application allows us to observe the evolution of the experiment, the results so far and with which activity it is being confused in case of a failure. All this at *interactive time* (also known as *operational time*) while the participants performed the activities. When we started to see that there were many mistakes, we started to pay attention to the predicted activity in the failure cases. We wanted to know if the confusion was due to any of the sink activities in particular, and if maybe there should be a change in the value of the limit selected.

But it was not like that. Most of the confusions in the detection of *Leg Flexion* and *Leg Circular Swing* happened with *Grab Object*, bumping down both activities to the second place of most similarity obtained (in most cases by a negligible difference in percentage, less than 0.2%). Therefore, the limits were working correctly, avoiding that sink activities would be the prediction even when obtaining a higher similarity percentage. But we did not think this could happen.

Seeing the results of *Grab Object* in the evaluations of the first and second strategy, it is obvious that it is an activity that is easier to detect than those performed with the legs (once the strategies proposed in this paper have been applied). But it is not an activity that we currently consider a sink activity. It is true that it fits the criteria of not being done suddenly. But on the other hand, its number of relevant joints is quite small, and it is an activity very prone to having postural coincidences that negatively affect its results.

Based on what happened, we can conclude that the strategy of applying limits to sink activities is an interesting idea. An idea that can contribute to improve recognition, but it needs to be refined.

The explanation for what may have happened is as follows. When asking the participants to perform an activity, they are shown how to make it, but avoid conditioning its performance. The only stricter guideline is that they must follow the speed of the activity to perform. There can be differences regarding the master sample, but the speed cannot differ significantly (the management of these differences is part of the future work). The theory is that, when performing activities with the legs, the movements done by the arms of the participants unconsciously to maintain balance, may have been confused with the movement *Grab Object*. This attached to the oscillations in which Kinect frequently incurs in, may have been the reason.

Lastly, if we compare on a general level the performance of both strategies (*specialized body parts analysis* vs. *stricter restrictions for the most easily detectable activities*) undoubtedly the first of those is the winner. Thanks to the excellent results obtained through the set vC as well as the set vT. The application of the *minimum limits of prediction* (personalized to each sink activity) has demonstrated to be an interesting strategy. One that works in the sense of limiting the prediction of sink activities only when they are being done (without reducing its accuracy rate). But its application by itself cannot be considered as a solution to the problem we are trying to solve.

Given the virtues that applying the limits has demonstrated, it is a strategy that maybe, with a little bit more refinement (and with a detailed study of the exact reasons of what has happened with *Grab Object*), it can end up as a useful tool for activity recognition.

6. Strategy #3—Combination of Body Parts and Stricter Limits

6.1. Description of the Strategy

Once we applied the two previously proposed strategies, the next logical step would consist in combining both strategies. The *specialized body parts analysis*, through the set of joints vT, has proven to be a possible solution to the problem of detect *bounded* activities (Section 3). On the other hand, the use of *stricter restrictions* for the most easily detectable activities has not worked as we expected, (recognition rates doesn't improve enough) but it has positive aspects (the limits work properly). For this third strategy, we have combined both approaches, with the aim of strengthening the recognition algorithm (*E-BA-A*) to have the advantages of both.

6.2. Description of the Strategy Evaluation

The dynamic of the evaluation is the same as in the first strategy (Section 4.2). The difference is that, apart from the sets vB, vT and vC, the *minimum limits of prediction* assigned to the sink activities are also present, and can be seen in Table 1.

6.3. Evaluation Results

The interpretation of Figure 7 and its results is equivalent to the interpretation of the results from the first evaluation (Section 4.3). Suc indicates the number of times that the algorithm predicted correctly the movement performed by the participant. Vic indicates the number of times that that set of joints (vB, vT or vC) has obtained a similarity with the activity greater than the rest of considered sets. The only difference is that, in this occasion, each volunteer has performed each activity a total of 25 times instead of 50. This way the evaluation is more aggressive, as any mistake will penalize the final accuracy rate much more than if we have a larger number of instances.

Combination of Body Parts and Stricter Restrictions Evaluation – Strategy #3 – Test Results – Sensors 2017																				
vB	Drink [D]				Grab Object [D]				Leg Flexion [I]				Leg Circular Swing [I]				Total			
	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate
Subject #1	23	92 %	0	0 %	17	68 %	4	16 %	21	84 %	0	0 %	25	100 %	2	8 %	86	86 %	6	6 %
Subject #2	25	100 %	0	0 %	25	100 %	0	0 %	0	0 %	0	0 %	0	0 %	0	0 %	50	50 %	0	0 %
Subject #3	25	100 %	0	0 %	25	100 %	25	100 %	0	0 %	0	0 %	0	0 %	0	0 %	50	50 %	25	25 %
Subject #4	17	68 %	0	0 %	25	100 %	3	12 %	0	0 %	0	0 %	4	16 %	0	0 %	46	46 %	3	3 %
Subject #5	25	100 %	0	0 %	25	100 %	0	0 %	7	28 %	0	0 %	2	8 %	0	0 %	59	59 %	0	0 %
Total	115	92 %	0	0 %	117	93.6 %	32	25.6 %	28	22.4 %	0	0 %	31	24.8 %	2	1.6 %	291	58.2 %	34	6.8 %
vT	Drink [D]				Grab Object [D]				Leg Flexion [I]				Leg Circular Swing [I]				Total			
	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate
Subject #1	25	100 %	25	100 %	20	80 %	18	72 %	25	100 %	12	48 %	25	100 %	22	88 %	95	95 %	77	77 %
Subject #2	25	100 %	25	100 %	25	100 %	25	100 %	25	100 %	3	12 %	25	100 %	9	36 %	100	100 %	62	62 %
Subject #3	25	100 %	25	100 %	25	100 %	0	0 %	25	100 %	0	0 %	25	100 %	22	88 %	100	100 %	47	47 %
Subject #4	25	100 %	25	100 %	25	100 %	22	88 %	23	92 %	1	4 %	21	84 %	13	52 %	94	94 %	61	61 %
Subject #5	25	100 %	25	100 %	25	100 %	25	100 %	22	88 %	18	72 %	20	80 %	15	60 %	92	92 %	83	83 %
Total	125	100 %	125	100 %	120	96 %	90	72 %	120	96 %	34	27.2 %	116	92.8 %	81	64.8 %	481	96.2 %	330	66 %
vC	Drink [D]				Grab Object [D]				Leg Flexion [I]				Leg Circular Swing [I]				Total			
	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate	Suc	Rate	Vic	Rate
Subject #1	22	88 %	0	0 %	8	32 %	0	0 %	25	100 %	13	52 %	25	100 %	1	4 %	80	80 %	14	14 %
Subject #2	25	100 %	0	0 %	25	100 %	0	0 %	25	100 %	22	88 %	25	100 %	16	64 %	100	100 %	38	38 %
Subject #3	25	100 %	0	0 %	22	88 %	0	0 %	25	100 %	25	100 %	23	92 %	3	12 %	95	95 %	28	28 %
Subject #4	25	100 %	0	0 %	21	84 %	0	0 %	23	92 %	22	88 %	25	100 %	16	64 %	94	94 %	38	38 %
Subject #5	25	100 %	0	0 %	25	100 %	0	0 %	22	88 %	4	16 %	24	96 %	9	36 %	96	96 %	13	13 %
Total	122	97.6 %	0	0 %	101	80.8 %	0	0 %	120	96 %	86	68.8 %	122	97.6 %	45	36 %	465	93 %	131	26.2 %

Figure 7. Combination of Body Parts and Stricter Limits Evaluation Results.

6.4. Analysis of the Results and Comparative with the Results of the Previous Strategies

The obtained results are, in a way, aligned with those obtained during the *specialized body parts analysis evaluation* (Section 4.4). As it happened then, the set vB is able to obtain high rates of success but only in the recognition of the activities performed with the arms (92% for *Drink [D]* and 93.6% for *Grab Object [D]*). In the case of activities whose relevant joints are concentrated on the legs, thanks to the combined use of the *minimum limits of prediction*, the results have experimented an improvement (*Leg Flexion [I]* has gone from 0% to 22.4% and *Leg Circular Swing [I]* from 9% to 24.8%). But this improvement is not enough. The global accuracy percentage has also improved, going from 52.42% for the set vB to 58.2%. Hence, the option of combining the set vB with the use of the *minimum limits of prediction*, allows to improve the results of the set vB on its own. But it is not a viable option to obtain high accuracy rates (more details on why the limits have not improved more the results of the set vB in Section 5.6).

With the sets vC and vT the situation is also the same of what previously happened. The use of the set vC, in combination with the *minimum limits of prediction*, is recommended as it has obtained an overall success rate of 93%. The limits have contributed to improving its results, as in the original evaluation, the global accuracy rate was of 92.42%. But again, and as it happened in that evaluation, the drawback is that even though it is a very good rate, it is not the set that has achieved the best results. The best global success rate obtained during the evaluation has been 96.2%, corresponding to the set vT.

In this last case, the accuracy rate has been somewhat lower to the original (97.37%). But given that the evaluation has been more aggressive (as we had a smaller number of instances performed) this does not imply that the results from the set vT worsen when used in combination with the limits. Given the obtained results, we can conclude that the use of limits as a support tool contributes to the improvement of activity recognition.

7. Discussion

One of the limitations of this work is the size of the test population. In order to perform the evaluations, we had the collaboration of five participants. It is true that the number of participants is small, but on the other hand, the number of samples classified by the *E-BA-A* during the evaluations was huge. The five participants performed a total of more than 3500 samples of the considered activities in the three evaluations.

Besides, in a system with the characteristics of the *E-BA-A*, it is much more important to have many samples to classify, rather than many different participants; because these ones are not part of any training set. With a single sample of each activity, the system is classifying, with good results, thousands of samples performed by five different persons. This makes it not so critical to have a large test population. Furthermore, even though having only five participants, characteristic and notable differences have been observed between the different strategies applied.

Regarding the extrapolation of the results to other proposals, the results and clues obtained in this work could be extrapolated mainly to other systems based on the same kind of technology. This is due to the particularities of the skeletal tracking joint oriented devices. But there are some specific lessons learned that can be theoretically generalized. During the application of the first strategy, the results clearly indicate that the use of the *trunk* can favor a better recognition of the activities. These findings about the significance of each body part in the recognition can be useful for systems not necessarily based on skeletal tracking joint oriented devices.

8. General Conclusions and Future Work

Starting with the first strategy, its analysis results (Section 4.4) clearly indicate that the best option to analyze *bounded* activities is used, alongside the joints of the extremities, also the joints of the trunk. The trunk eases the obtention of high similarity percentages because its monitorization is more precise. It is a part of the body whose range of movements is more limited, facilitating its tracking with joint oriented devices.

The trunk provides stability to the results. It also reduces the factor that, to contemplate less *relevant bones*, they have more influence on the obtained similarities. And at the same time, they are more exposed to the sensor device possible failures. By including the trunk, we achieved that the activity has an *intermediate amount of relevant joints* (and therefore *relevant bones*), which ensures a balanced weight distribution in the obtaining similarities process.

But the use of the trunk alongside the extremities has not only been beneficial regarding the obtention of better rates of success. It is also useful to improve the victory percentage. Getting the set *vT* to obtain both a *greater overall success rate* (97.37%) as well as the *highest percentage of wins* (61.26%). As it is able to obtain higher similarities (than the rest of alternatives sets), potentially it will also obtain greater differences regarding the similarities obtained for the rest of the activities that have not been done by the user. As a result, obtaining greater differences with the rest of activities, theoretically eases the differentiation between activities.

Regarding the application of greater restrictions for the most easily detectable activities (second strategy), it has not worked as we expected. It is an interesting idea because the restrictions that we applied worked well. The limits were able to avoid the prediction of *sink* activities when they were not being performed. But, by itself, it is not a solution to improve the recognition of *bounded* activities.

But on the other hand, the results of the third strategy, indicates us that the use of the limits could be useful as a support tool. The addition of the limits, in combination with the first strategy, improves the recognition rates of the sets *vB* and *vC* and the success percentages of the set *vT* (this last one going from 61.26% to 66%).

The results obtained in this work pretend to be some first indications about what considered strategy is better. Taking into account that the population of volunteers that we have used is small, we offer some preliminary clues for other researchers who are working on systems based on the same kind of technology.

Apart from increasing the size of the test population, future work goes through including distances into the *E-BA-A*. This way we will be able to improve the results obtained by this last one enriching the information it works with. Going from only using the angles to detect the activities, to combining the angles with the distances the joints have moved. This measure could be the solution to the conflict experienced with the activity *Grab Object* (Section 5.6) as to reach a decision, the distance moved by the leg joints would also be taken into account.

Other lines of future work consist in maximizing the use of the algorithm collaborating with other authors. The idea is to add value to other works. On the one hand we will combine this system with an affective avatar [28] in order to give this last one information about the activities that the user is doing. On the other hand, we will develop a videogame whose main mechanic will be the realization of physical activities. This will allow to analyze the relationship between the game mechanic and the executive function [29]. The activities performed during the game will be analyzed through the *E-BA-A*.

Author Contributions: C.G.-L.-F. and E.J. wrote the paper; C.G.-L.-F. designed and implemented all the software required to performed the experiments; C.G.-L.-F. and R.H. conceived and designed the experiments; C.G.-L.-F. and E.J. performed the experiments; C.G.-L.-F. and R.H. analyzed the results of the evaluations; C.G.-L.-F. and R.H. studied the state of the art; R.H. made the final revision of contents; E.J. translated the full article.

Funding: This work was conducted in the context of UBIHEALTH project under International Research Staff Exchange Schema (MC-IRSES 316337) and the coordinated project grant TIN2013-47152-C3-1-R (FRASE), funded by the Spanish Ministerio de Ciencia e Innovación.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A. Activities Considered during the Performance of the Evaluations

Each of the evaluations recorded in this paper has as the object of study a group of activities that is not the same in all of the cases. Also, none of the cases in that group of activities coincides completely with the contemplated activities in the system's data base. Therefore, even though we evaluate the development of each activity that is of interest in each evaluation, this does not mean that they are the only activities that are considered by the *E-BA-A*. The goal is to test if the proposed solutions allows to improve the recognition of activities with a more reduced subset of relevant joints. Hence, it is necessary that the algorithm contemplates more activities in order to make the task more complicated. The total of activities considered by the *E-BA-A* in all of the performed tests can be seen in Table A1 (which contemplates the characteristics previously explained about *Symmetry* and *Global/Bounded Use of the Body*).

Table A1. Taxonomy of the considered activities by the *E-BA-A* during the tests.

	Symmetric	Asymmetric
Global	Forward Bend	
	Sit Down	Walk
	Stand Up	Walk Backwards
	Stand	
	Stay Seated	
Bounded		Drink
	Squat	Grab Object
		Leg Flexion
		Leg Circular Swing

It is worth mentioning the case of the activities *Walk* and *Walk Backwards*. As it can be seen in the evaluation corresponding to the second strategy considered for this work (Section 5), these activities are signaled with a $[D]$. The reason is that the samples were recorded starting the walk with the right

leg. This detail is indifferent as long as the user that is being monitored walks more than one stride cycle. But if only one is made, the leg with which the user starts to walk with has to be taken into account. Because of this, in the future, we will also include in the system's data base a sample of the activities *Walk [I]* and *Walk Backwards [I]*.

Appendix B. Body Information of the Participants of the Evaluations

The evaluations have been done thanks to the participation of five volunteers of different ages, genders (two men and three women) and complexions. These last are detailed in Table A2. The participants have no disabilities nor specialized background in any athletic discipline.

Table A2. Test Subjects Corporal Information.

Test Subject	Height	Shoulder Width (cm)	Hip Width (cm)	Age
#1	1.60	39	45	26
#2	1.81	50	42	57
#3	1.56	44	38	56
#4	1.68	37	34	23
#5	1.84	45	44	26

The samples used by the *E-BA-A* (only one for each considered activity) for training were performed only by the test subject #5. This means that we tested the performances of the subjects #1, #2, #3 and #4 with the information of only one person (the test subject #5).

Appendix C. Hardware Used During the Evaluations

The hardware used during the evaluation was a standard desktop computer whose main components correspond with a RAM of 8 GB DDR4 and an Intel Quad-Core i5-6500HQ processor operating at 2.3 GHz. The *Extended Body-Angles Algorithm* not requires specialized hardware and even it is capable of running on older hardware configurations, as it is explained in greater detail in previous publications [5]. Depending on the performance of the used equipment, the main difference is the sampling rate with which the sensor device is capable of perceive its surroundings. And in consequence, the samples that the algorithm processes per second is different too.

Appendix D. Particularities of the Evaluations and Data Gathering Protocol

In this work, we have used an original method, named *E-BA-A*, in order to do the classification of the activities, whose successful performance was validated in previous works [5]. At first sight, this method could be identified as just another proposal based on machine learning techniques. However, and considering the Davis and Marcus taxonomy [6], this kind of interaction cannot be considered machine learning. Actually, we are proposing a commonsense reasoning method based on mathematical analogies for activity recognition. For this reason, it is convenient to clarify some aspects which sometimes can be confusing. These two aspects, related with the development of the evaluations, are: the fact that *there is no need to record* the performed activities and the *simultaneous obtention* of the results.

In more detail, for a better understanding of this recognition algorithm, it is important to remark that the *E-BA-A* do not need to record large data corpora to obtain the representation of the activities. In consequence, we do not create a dataset with multiple samples of each activity in order to process that information later. The *E-BA-A* is capable of classifying activities using only one sample of each one through mathematical analogies. This means that, when the participants performed the activities, all the instances are directly classified by the *E-BA-A*. The thousands of instances of the considered activities performed by the participants (more than 3500 taking into account all the evaluations) are only for test purposes, not for training.

On the other hand, and taking the previous point into account, we obtain the results of the evaluations in a simultaneous way, during its realization. The *E-BA-A* works at interactive time (also known as operational time). This means that, the results of the evaluation and the predictions of the *E-BA-A*, are obtained at the same moment that the participant performs the activity.

Once we explained these particularities, we are going to detail the *data gathering protocol*. As the *E-BA-A* is not a system based on machine learning, the concepts of *training set* and *test set* are a little bit blurred. The reason for that, as we previously introduced in the paper, is that the *E-BA-A* only needs one sample of each considered activity (training set). In consequence, all the activities performed during the evaluations are for test purposes only (test set). But we are going to use these concepts to facilitate the comprehension of the data gathering. In order to build the training set (only one sample of each activity), before the beginning of the evaluation, we asked one of the participants to perform every considered activity only one time. We recorded this performance in the database of the *E-BA-A* using *Microsoft Kinect* (specifically its skeleton stream). Each element of the database contains the position of the joints of the body in each frame of the activity. This way, the *E-BA-A* will use these unique samples of each activity to do the classification of future samples.

Regarding the building of the test set, during the evaluations, we asked the participants to perform the repetitions of the activities in front of Kinect. But these repetitions were not recorded. This means that, all of these repetitions are used as a test set, but are used on the fly. Thanks to the capabilities of the *E-BA-A*, in the exact moment in which the participant finishes the performing of every repetition, we obtained the classification result of that repetition.

References

1. Gutiérrez López de la Franca, C.; Hervás, R.; Johnson, E.; Bravo, J. Findings about Selecting Body Parts to Analyze Human Activities through Skeletal Tracking Joint Oriented Devices. In Proceedings of the 10th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2016), Gran Canaria, Spain, 29 November–2 December 2016; García, C.R., Caballero-Gil, P., Burmester, M., Quesada-Arencibia, A., Eds.; Springer: Berlin, Germany, 2016; pp. 537–548.
2. Fathi, A.; Farhadi, A.; Rehg, J.M. Understanding egocentric activities. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011; pp. 407–414.
3. Cleland, I.; Kikhia, B.; Nugent, C.; Boytsov, A.; Hallberg, J.; Synnes, K.; Finlay, D. Optimal placement of accelerometers for the detection of everyday activities. *Sensors* **2013**, *13*, 9183–9200. [[CrossRef](#)] [[PubMed](#)]
4. Kusumam, K. Relational Learning Using Body Parts for Human Activity Recognition Videos. Master's Thesis, University of Leeds, Leeds, UK, 2012.
5. Gutiérrez López de la Franca, C.; Hervás, R.; Johnson, E.; Mondéjar, T.; Bravo, J. Extended Body-Angles Algorithm to Recognize Activities within Intelligent Environments. *J. Ambient Intell. Hum. Comput.* **2017**, *8*, 531–549. [[CrossRef](#)]
6. Davis, E.; Marcus, G. Commonsense reasoning and commonsense knowledge in artificial intelligence. *Commun. ACM* **2015**, *58*, 92–103. [[CrossRef](#)]
7. Alexiadis, D.S.; Kelly, P.; Daras, P.; O'Connor, N.E.; Boubekour, T.; Moussa, M.B. Evaluating a dancer's performance using kinect-based skeleton tracking. In Proceedings of the 19th ACM international conference on Multimedia, Scottsdale, AZ, USA, 28 November–1 December 2011; pp. 659–662.
8. Leightley, D.; Darby, J.; Li, B.; McPhee, J.S.; Yap, M.H. Human Activity Recognition for Physical Rehabilitation. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Manchester, UK, 13–16 October 2013; pp. 261–266.
9. Raso, I.; Hervás, R.; Bravo, J. m-Physio: Personalized accelerometer-based physical rehabilitation platform. In Proceedings of the Fourth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, Florence, Italy, 25–30 October 2010; pp. 416–421.
10. Fontecha, J.; Navarro, F.J.; Hervás, R.; Bravo, J. Elderly frailty detection by using accelerometer-enabled smartphones and clinical information records. *Pers. Ubiquitous Comput. J.* **2013**, *17*, 1073–1083. [[CrossRef](#)]

11. Goncalves, N.; Costa, S.; Rodrigues, J.; Soares, F. Detection of stereotyped hand flapping movements in Autistic children using the Kinect sensor: A case study. In Proceedings of the IEEE Autonomous Robot Systems and Competitions (ICARSC), Espinho, Portugal, 14–15 May 2014; pp. 212–216.
12. Vrigkas, M.; Nikou, C.; Kakadiaris, I.A. A Review of Human Activity Recognition Methods. *J. Front. Robot. AI* **2015**, *2*, 28. [[CrossRef](#)]
13. Brush, A.; Krumm, J.; Scott, J. Activity recognition research: The good, the bad, and the future. In Proceedings of the Pervasive Workshop How to do Good Research in Activity Recognition, Helsinki, Finland, 17–20 May 2010.
14. Veltmaat, M.J.T.; van Otterlo, M.; Vogt, J. *Recognizing Activities with the Kinect*; Technical Report; Radboud University: Nijmegen, The Netherlands, 2013.
15. Ladha, C.; Hammerla, N.; Hughes, E.; Olivier, P.; Plötz, T. Dog's life: Wearable activity recognition for dogs. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, Zurich, Switzerland, 8–12 September 2013; pp. 415–418.
16. Fu, J.; Liu, C.; Hsu, Y.P.; Fu, L.C. Recognizing context-aware activities of daily living using RGBD sensor. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Tokyo, Japan, 3–7 November 2013; pp. 2222–2227.
17. Tapia, E.M.; Intille, S.S.; Larson, K. Activity recognition in the home using simple and ubiquitous sensors. In Proceedings of the International Conference on Pervasive Computing, Linz and Vienna, Austria, 21–23 April 2004; Ferscha, A., Mattern, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3001, pp. 158–175.
18. Ryoo, M.S.; Matthies, L. First-Person Activity Recognition: What Are They Doing to Me? In Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR), Portland, OR, USA, 2013; pp. 2730–2737.
19. Yang, Z.; Zicheng, L.; Hong, C. RGB-Depth feature for 3D human activity recognition. *Commun. J.* **2013**, *10*, 93–103.
20. Biswas, K.K.; Basu, S.K. Gesture Recognition using Microsoft Kinect®. In Proceedings of the 5th International Conference on Automation, Robotics and Applications (ICARA), Wellington, New Zealand, 6–8 December 2011; pp. 100–103.
21. Ramirez-Giraldo, D.; Molina-Giraldo, S.; Alvarez-Meza, A.M.; Daza-Santacoloma, G.; Castellanos-Dominguez, G. Kernel based hand gesture recognition using kinect sensor. In Proceedings of the IEEE XVII Symposium on Image, Signal Processing, and Artificial Vision (STSIVA), Antioquia, Turkey, 12–14 September 2012; pp. 158–161.
22. Boger, J.; Hoey, J.; Poupart, P.; Boutilier, C.; Fernie, G.; Mihailidis, A. A planning system based on Markov decision processes to guide people with dementia through activities of daily living. *IEEE Trans. Inf. Technol. Biomed. J.* **2006**, *10*, 323–333. [[CrossRef](#)]
23. Bansal, S.; Khandelwal, S.; Gupta, S.; Goyal, D. Kitchen activity recognition based on scene context. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Melbourne, Australia, 15–18 September 2013; pp. 3461–3465.
24. Banos, O.; Toth, M.A.; Damas, M.; Pomares, H.; Rojas, I. Dealing with the effects of sensor displacement in wearable activity recognition. *Sensors* **2014**, *14*, 9995–10023. [[CrossRef](#)] [[PubMed](#)]
25. Kunze, K.; Lukowicz, P. Sensor placement variations in wearable activity recognition. *IEEE Perv. Comput. J.* **2014**, *13*, 32–41. [[CrossRef](#)]
26. Kunze, K.; Lukowicz, P.; Junker, H.; Tröster, G. Where am I: Recognizing on-body positions of wearable sensors. In Proceedings of the First International Symposium on Location and Context-Awareness (LoCA), Oberpfaffenhofen, Germany, 12–13 May 2005; pp. 264–275.
27. Polana, R.; Nelson, R. Low level recognition of human motion (or how to get your man without finding his body parts). In Proceedings of the IEEE Workshop on Motion of Non-Rigid and Articulated Objects, Austin, TX, USA, 11–12 November 1994; pp. 77–82.

28. Johnson, E.; Hervás, R.; Gutiérrez López de la Franca, C.; Mondéjar, T.F.; Ochoa, S.; Favela, J. Assessing Empathy and Managing Emotions Through Interactions with an Affective Avatar. *Health Inf. J.* **2016**. [[CrossRef](#)] [[PubMed](#)]
29. Mondéjar, T.; Hervás, R.; Johnson, E.; Gutiérrez, C.; Latorre, J.M. Correlation between videogame mechanics and executive functions through EEG analysis. *J. Biomed. Inf.* **2016**, *63*, 131–140. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).