Preliminary study in healthy subjects of arm movement speed assessment by Kinect

Mohamed Elgendi[1,*], Flavien Picon[2], Nadia Magnenat-Thalmann[2], Derek Abbott[3]

1 Department of Computing Science, University of Alberta, Canada

2 Institute of Media Innovation, Nanyang Technological University, Singapore

3 School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, South Australia, Australia

∗ **E-mail: moe.elgendi@gmail.com**

Abstract

Many clinical studies have shown that the arm movement of patients with neurological injury is often slow. In this paper, the speed analysis of arm movement is presented, with the aim of evaluating arm movement automatically using a Kinect camera. The consideration of arm movement appears trivial at first glance, but in reality it is a very complex neural and biomechanical process that can potentially be used for detecting a neurological disorder. This is a preliminary study, on healthy subjects, which investigates three different arm-movement speeds: fast, medium and slow. With a sample size of 27 subjects, our developed algorithm is able to classify the three different speed classes (slow, normal, and fast) with overall error of 5.43% for interclass speed classification and 0.49% for intraclass classification. This is the first step towards enabling future studies that investigate abnormality in arm movement, via use of a Kinect camera.

Introduction

Slowness in arm movement is common in many disorders, such as Huntington's chorea [1], Parkinson's disease [2] and cerebellar diseases [3]. However the abnormality in arm movement varies from one disease to another. Given the vast array of disorders associated with abnormal movements, the challenge for the rehabilitation community is in obtaining high quality evaluations at low cost.

Recently, Kinect cameras offer extremely inexpensive accurate information for sensitive motion tracking [4]. Moreover, Kinect is also considered as a promising tool for the investigation of tremor and

slowness in arm movement [5]. To our knowledge, there are no studies that investigate the speed of arm-movement joints for detecting abnormality in arm movements using Kinect or any other depth cameras. However, several arm-movement recognition systems have considered the speed as a feature. Min et al. [6] confirmed that arm-movement recognition is usually dependent on the trajectory of arm movement, and that position, speed, and curvature are useful features. Campbell et al. [7] investigated ten different features for arm-movement recognition using a Hidden Markov Model (HMM). They indicated that speed features are superior to positional features. Yoon et al. [8] used hand speed as an important feature for arm-movement recognition.

Other researchers estimated the speed of arm-movement using an accelerometer, for example Rehm et al. [9] used the power of the accelerometer as a feature to classify the arm movement into low and high speed regimes, not for diagnosis purposes. However, a recent clinical study by Howard et al [10] explored spatial and temporal changes in shoulder motion, in both healthy asymptomatic healthy adults and rotator cuff patients, during different speeds of movement using a scapular tracking device. In contrast to those studies, in this paper we systematically explore the speed of arm-movement joints, with the aim of improving the classification of the arm-movement speed. Our study builds upon Rehm's and Howard's work by providing a device-free analysis of arm movement, exploring the impact of different joints on the overall arm movement, and validating the system in a noisy environment.

Materials and Methods

Ethics Statement

In this study, no film recordings of subjects were made. The Kinect camera outputs numerical data that directly relate to hand movements. Only de-identified numerical data, representing motion vectors, are stored on the database. Volunteers are researchers at office of the Institute for Media Innovation, NTU, Singapore. All data is available at:

`http://www3.ntu.edu.sg/imi/piconflavien/autres/data-speed-arm.zip` and

`http://www.elgendi.net/databases.htm` .

Data Collection

There are currently no standard Kinect databases for arm-movement analysis available to evaluate our developed algorithm. However, the Institute of Media Innovation at Nanyang Technological University has one database that contains arm movement data of 27 healthy volunteers (6 females and 21 males); with a mean ± SD age of 29.7 ± 4.1, height of 172.9 cm ± 9.3 cm, arm length of 71.3 cm ± 5.2 cm. Two of them were left-handed. The motions were measured using the Kinect camera located 2.7 meters away from the subject at a height of 1.2 meters above the floor, cf. Figure 1. All Kinect data is acquired using the Microsoft Kinect SDK Beta 1 (Microsoft, 2012) at a sampling frequency of 30 Hz. The Kinect device consists of laser light source, color camera and an infrared camera. The infrared laser source and the infrared video camera form the depth camera function, while the colour video camera provides colour data to the depth map. The technology was developed by PrimeSense (Tel-Aviv, Israel) and is disclosed in detail in their patents [11].

During the experiment, the body of the subject faces the sensor with an angle of 45° to the right of the Kinect sensor (as seen in Figure 1). The reason behind the 45° angle is to prevent the arm joints from intersecting with the body joints, as shown in Figure 2. This will generate reliable arm motion in order to study the impact of each joint of the arm on the overall speed of the right arm movement more precisely. These collected arm movements are used as a benchmark for effective speed detection of an arm movement. Measurements were taken with each subject standing vertical, with an initial position where both arms extended along the body side. Then, the subject is asked to raise a right arm up. Each subject performs three sets of trials: 'slow', 'normal', and 'fast'; with five arm movements for each set. Therefore the number of recorded movements is 405 (27 subjects × 5 movements × 3 speeds).

For the slow movement, the subject is instructed to raise an arm as if there is a heavy weight being lifted to simulate a typical indicator of Bradykinesia (a symptom of nervous system disorders, particularly Parkinson's disease). On the other hand, the fast arm movement indicates healthy motion, while the normal speed represents an average condition. Capture of the arm movement is carried out manually. In other words, the subjects wait for a signal from the recording person to start their movement and then they maintain their arm up until they get a signal to come back to the initial position. Each recording is played back, checked, and annotated as being in one of three classes 'slow', 'normal', or 'fast'. Two independent annotators annotated the speed category of each recorded movement; when two annotators disagree, the result is discarded and the subject is asked to repeat the experiment. The annotations were

stored in a file to be compared automatically later with the speed features that will be discussed in the next section.

Methodology

The proposed arm-movement classification type algorithm consists of three main stages: pre-processing (resultant of coordinates as instantaneous velocity and low-pass filtering), feature extraction (calculating the first and second derivative and their mean and standard deviation) and classification (thresholding). The structure of the algorithm is shown in Figure 3.

Pre-Processing

The Kinect body tracking software API provides the real-time position of the body joints of each user [12]. Even though we focus mainly on the skeletal joints of the arm we chose to record the positions of all skeletal joints: center of gravity or legs movements are also potential speed indicators. With 20 joints and 3 floating point values (real numbers) representing the x, y, z positions for each joint, each motion frame is expressed as a 60-element vector. The recorded joints cover all parts of the body but we focus mainly on the arm joints: shoulder, elbow, wrist and hand. Since the features only rely on the dynamics of the motion there are no differences in processing data from the left or right arm. Therefore we may process data from the joints of each subject's dominant arm (25 right-handed and 2 left-handed).

The 3D position (x, y, z) of a joint is expressed in the coordinate system of the Kinect and the units are meters [13]. Again, the selected features rely on motion dynamics so our system is view-independent: we do not have to express the positions in the coordinate system of the subject's body. The dynamic of each joint is computed using the variation of position of the joint over time. In the first step, each joint motion, sequence of 3D positions, is replaced by the distance between each frame as in Eq. 1. In Figure 2, the x, y, z coordinates are the positions vectors of a particular joint that vary 0 to n; where n is the number of frames in a performed motion. The instantaneous velocity of motion for a particular joint are calculated as the resultant of x, y, z positions over all frames that represents a motion. The instantaneous velocity (u_{inst}) for a given 3D motion is computed as follows:

$$u_{\text{inst}}[n] = \left.\frac{dx, y, z}{dt}\right|_{t=nT} = \frac{1}{T}\sqrt{(x[n]-x[n-1])^2 + (y[n]-y[n-1])^2 + (z[n]-z[n-1])^2}, \qquad (1)$$

where T is the sampling interval and equals the reciprocal of the sampling frequency, and n is the number of motion data points.

As shown in Figure 4, the informative part of the motion lies below 6 Hz for all joints with different speed types. Thus, a low-pass filter has been applied. A first order, zero-phase bidirectional, Butterworth low-pass filter with cutoff frequency of 6 Hz is implemented. Figure 5, shows an example of the original data u_{inst}, at top-left, and the filtered data (V_{inst}), at top-right, with no phase distortion. Note that the low frequencies will play a major role in identifying hand movement speed, and ultimately hand tremors. The first order filter has been selected to avoid over-smoothing the acquired motion. This has been carried out empirically to find a condition where the substantial part of the motion is preserved while sensor errors were strongly reduced. We decided to record the raw data, i.e. without using the pre-defined filter provided in the Kinect SDK. By doing so, we have more control over the data analysis. We then have freedom to examine the affect of filtering on the classification rate.

Feature Extraction

Before continuing the discussion of the joint signals, it is important to know what features can be extracted from a hand movement first. In the literature, the instantaneous velocity and acceleration have been used in diagnosing arm movements. Almeida et al. [14] examined individuals with Parkinson's disease through the analysis of the upper-limb movement at different movement frequencies, and with different external timing conditions using the instantaneous velocity. However, investigators [15, 16] used the instantaneous velocity and acceleration to investigate the movements of the finger, elbow, and shoulder during a speed aiming movements. In this paper, two features are investigated: the instantaneous velocity and acceleration. The mathematical definition of the instantaneous velocity (u_{inst}) in 3D motion before filtering is described in Eq. 1, while the instantaneous acceleration (A_{inst}) is defined as:

$$A_{\text{inst}}[n] = \left.\frac{dV_{\text{inst}}}{dt}\right|_{t=nT} = \frac{1}{T}(V_{\text{inst}}[n] - V_{\text{inst}}[n-1]). \tag{2}$$

Although the Kinect camera has received increasing attention recently, it nevertheless suffers from noise, low resolution sensors, lack of color information, and occlusion problems [17]. Therefore, it is crucial that we filter the signal to improve the classification accuracy, especially if the main goal is to determine the speed type. In our study, we computed the instantaneous velocity and instantaneous acceleration

for each arm joint. Then, we calculated the following measures: average (f_1, f_3) and standard deviation (f_2, f_4). Two features $\{f_1, f_2\}$ extracted from the velocity V_{inst} and two features $\{f_3, f_4\}$ extracted from the acceleration A_{inst}. Features $f_1, f_2, f_3, \text{and} f_4$ are calculated as follows:

$$f_1 = (1/N) \sum_{n=1}^{N} V_{\text{inst}}[n], \tag{3}$$

$$f_2 = \sqrt{(1/N) \sum_{n=1}^{N} (V_{\text{inst}}[n] - f_1)^2}, \tag{4}$$

$$f_3 = (1/N) \sum_{n=1}^{N} A_{\text{inst}}[n], \tag{5}$$

$$f_4 = \sqrt{(1/N) \sum_{n=1}^{N} (A_{\text{inst}}[n] - f_3)^2}, \tag{6}$$

where N refers to the total number of samples in the processed motion.

Figure 5 demonstrates the signal shape of four different joints of a arm movement based on the instantaneous velocity and acceleration. This is particularly interesting as it confirms that joints of a same limb have the same dynamics, especially for the hand and wrist signals. As the variance of the hand and wrist joint signals are quite higher compared to the elbow and shoulder signals, it is expected that the hand or the wrist signal would score higher accuracy in the classification of arm movements.

Classification

In this section, we check the linear separability of the calculated feature set $f = \{f_1, f_2, f_3, f_4\}$ in both filtered and non-filtered signals. The classification steps are:

1. exclude one subject for testing and the rest of the subjects are used for training, following the leave-one-out (LOO) cross-validation process.
2. sort the motions
   (a) Inter: for the **training** subjects, 390 motions in total, based on the selected feature. We separate the motions in three clusters based on the feature value;
   (b) Intra: for **each** subjects of the **training** dataset, 15 per subject, based on the selected feature.

For **each subject** of the **training** dataset we separate the motions in three clusters based on the feature value;

3. compute the classification accuracy by counting the number of misclassified motions over the total number of motions of the 26 subjects (after excluding one subject during the LOO operation). A misclassified motion is a motion which feature based class is different from the recording's class;
   (a) Inter: done over all the training dataset.
   (b) Intra: done independently for **each** subject in the training dataset.
4. compute the two thresholds used to separate the three classes.

$$\text{slow} \implies \begin{cases} \quad \text{if } f(i) < \text{THR}_1(i), \end{cases}$$

$$\text{medium} \implies \begin{cases} \quad \text{if } \text{THR}_1(i) < f(i) < \text{THR}_2(i), \end{cases}$$

$$\text{fast} \implies \begin{cases} \quad \text{if } f(i) > \text{THR}_2(i), \end{cases}$$

   where i refers to the feature used $f_1, f_2, f_3, \text{or} f_4$.

   They represent the two values that minimise the difference between feature classification and recording classification;

   (a) Inter: done over all the training dataset ;
   (b) Intra: done independently for **each** subject in the training dataset. Averaging the two thresholds over **all subjects** of the **training** dataset to get two final thresholds for further classification use;
5. compute the classification error on the **test** dataset using the thresholds from the previous step;
6. perform the previous steps (2 to 7) using each subject as training data. A total of 27 thresholds computations and evaluations will be performed.

For each subject, used as a test dataset, the thresholds and the error are reported, cf. table2. In this table only the results from the most relevant features are selected: mean and standard deviation of instantaneous velocity of the hand $\{f_1, f_2\}$. The results in the table show that the thresholds are quite similar for the different training datasets. We can also notice that there is some subjects for whom

the classification error is pretty high. These errors are come from the strict separation provided by the thresholds. The first classifier is fast/medium against slow (THR_1), while the second classifier is fast against medium/slow (THR_2).

Figure 6 demonstrates the thresholds determination for inter- and intra- class speed classification. The two valleys reflect the thresholds that will be used for training the automatic speed detection. For example, in case of intra- class speed classification, the slow-medium threshold is 0.58 while the medium-fast threshold is 1.50 in non-filtered condition as shown in Figure 6 (left).

Results and Discussion

Advances in microelectromechanical systems allow measurement of the changes in velocity, position and acceleration by enabling low-cost sensors, accelerometers and gyroscopes. These sensors have been employed to analyze arm movement disorder.

In the literature, essential tremor typically has a frequency of 4–8 Hz when it is assessed by microelectromechanical systems, such as accelerometer [18]; however, our analysis shows frequencies that lie before 6 Hz are more informative for arm-speed assessment using Kinect. Certainly we need tremor measurements to draw a subtle conclusion here.

The statistical Kruskal-Wallis and ANOVA tests allow us to investigate whether the hand-movement speed feature takes different values among the three speed classes. In case of interclass analysis for non-filtered/filtered f_1, $p < 0.00001$ is scored by the ANOVA test, while the the Kruskal-Wallis test showed significance with $p < 0.01$. On the other hand, the intraclass analysis for filtered f_2, $p < 0.000001$ is scored by the ANOVA test, while the Kruskal-Wallis test showed significance with $p < 0.01$.

In the case of interclass speed analysis, low p-values ($p < 0.00000001$) have been scored for both tests, which indicate large difference in the means and medians of the three speed classes. Both tests find that the three hand-movement speeds are significantly different in the case of filtered and non-filtered features. The very small p-value indicates that differences between the three speed classes are highly significant.

In Table 1, as expected, the hand joint was successfully classified into different speed types with the lowest error rate (0.49% for intra classification and 5.43% for inter classification). This result confirms the observation, shown in Figure 5, which is that the the mean of the instantaneous velocity for the hand motion contains more information compared to the other three joints in both cases filtered and un-filtered

data. As can be seen, the hand joint is the most reliable for detecting speed in arm movement. It is interesting to note that features based on standard deviation perform better than those based on mean. Interestingly, the results of filtered and non-filtered hand-joint signal are relatively close. However, the filtered hand-joint signal scored a slightly lower classification error compared to the non-filtered signal.

Is it possible to predict the speed type before the completion of a full hand movement? To answer this, we investigated the percentage of the hand-movement from the start of the motion that contributes the most to the classification error. The results of this investigation are shown in Figure 7. For the inter-classification, the first 50% of the hand instantaneous velocity provides classification error 8.8% for SD feature. While in case of intra-class speed classification, the error rate scored by the first 60% was 4.4% for the mean without filtering. This is an interesting observation as the first 50% of a motion provides low classification error and relatively close in terms of performance to the whole motion. Knowing this fact can lead to an effective prediction, which can be even done in realtime without waiting for the whole motion to be completed. What portion of a hand-movement signal contributes the most to the classification error? Which 10% portion of the motion's signal contains the most useful information to distinguish the speed types?

Figure 8 shows the error rate for a sequential 10% of the motion signals. It can be seen, in the case of inter-class speed classification; the portion 50–60% of the mean of the hand instantaneous velocity provides the lowest error rate of 6.4%. This is intuitive as the beginning of motion is the phase where the subjects is reaching a certain poses. Moreover, the intraclass analysis shows that a 0% error rate can be achieved if the 50–60% portion of the mean of the instantaneous velocity is used, instead of the whole hand movement. This confirms that the main characteristics of a motion is determined within 50–60%, and can be used for analysis or/and prediction.

Limitations of Study and Future Work

We recommend that future work examine our method on patients that suffer from hand tremor as the thresholds calculated in this study are based on healthy subjects. However, mimicking unhealthy motion provides intial values for the system before assessing real patients at a hospital/clinic. A larger sample size and a diverse set of tremor movements are needed in order to generalize the findings of this study. To our knowledge, there is no available Kinect database measured from patients with hand tremor.

In future studies it may be advisable to test the optimal distance for positioning Kinect as sometimes subjects cannot be detected if they are relatively close to the camera. It would be also useful to know how accurately the Kinect can estimate speed of arm movement compared the speed of arm movement with a benchmark standard (such as three dimensional analysis system, e.g. Vicon, Optitrak, etc.).

Perhaps combining all the features together to build a single classifier to separate the 3 classes of movement speed is one of our next steps.

Figure 9 shows that the sampling rate of Kinect camera is relatively unstable. It fluctuates between 26.95 Hz and 33.67 Hz, rather than sampling the data with consistent frequency at 30 Hz. Our results show that this is not an issue unless high frequency analysis is needed, which is not the main concern of our study. However, the use of smoothing techniques decreases the impact of the sensor instability.

Technically, exploring simple features such as the mean and standard deviation of a motion is promising in terms of computational complexity and efficiency. However, this can be further improved by investigating other features in time and frequency domain.

Thresholds provide a strict separation between the data, the resulting misclassification could be reduced by introducing an acceptance interval around the threshold values.

Conclusion

In this paper we presented a speed analysis of arm movement. Results show that: 1) the instantaneous velocity provides more reliable classification compared to the instantaneous acceleration, 2) the mean is a better feature compared to the standard deviation for the instantaneous velocity, and 3) the hand joint is the most efficient joint for speed detection in an arm motion. Moreover, a low-pass filter improves the interclass speed classification but has no affect on the intraclass classification. For interclass speed classification, the mean of non-filtered instantaneous velocity scored a 0.49% error rate in detecting the speed type over 405 motions; while the standard deviation of filtered instantaneous velocity scored 5.43% during the intraclass classification. Moreover, the first 60% provides a classification error relatively close to the use of a whole motion, can be used for predicting the speed type in realtime. Furthermore, the most important 10% of a whole motion is the 50–60% region. The results are promising and this approach can be implemented in a human-computer-intraction system for intractive tremor diagnosis, specifically measuring hand-related disability and improvement. In our approach we asked healthy subjects to mimic

abnormality by moving slowly, however testing this approach on patients with Parkinson's disease or any hand tremors remains a task for future work.

Acknowledgments

This research, which is carried out at BeingThere Centre, is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

References

1. Bittenbender J, Quadfasel F (1962) Rigid and akinetic forms of Huntington's chorea. Archives of Neurology 7: 275-288.

2. Jahanshahi M, Jenkins IH, Brown RG, Marsden CD, Passingham RE, et al. (1995) Self-initiated versus externally triggered movements I. an investigation using measurement of regional cerebral blood flow with pet and movement-related potentials in normal and Parkinson's disease subjects. Brain 118: 913–933.

3. Deuschl G, Wenzelburger R, Löffler K, Raethjen J, Stolze H (2000) Essential tremor and cerebellar dysfunction clinical and kinematic analysis of intention tremor. Brain 123: 1568–1580.

4. Weber I, Koch J, Meskemper J, Friedl K, Heinrich K, et al. (2012) Is the MS Kinect suitable for motion analysis? Biomedical Engineering / Biomedizinische Technik 57: 664.

5. Nguyen HA, Auvinet E, Mignotte M, de Guise JA, Meunier J (2012) Analyzing gait pathologies using a depth camera. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 4835–4838.

6. Min BW, Yoon HS, Soh J, Yang YM, Ejima T (1997) Hand gesture recognition using hidden Markov models. In: 1997 IEEE International Conference on Systems, Man, and Cybernetics. volume 5, pp. 4232–4235.

7. Campbell LW, Becker DA, Azarbayejani A, Bobick AF, Pentland A (1996) Invariant features for 3-D gesture recognition. In: Proceedings of the Second International Conference on Automatic Face and Gesture Recognition. pp. 157–162.

8. Yoon HS, Soh J, Bae YJ, Yang HS (2001) Hand gesture recognition using combined features of location, angle and velocity. Pattern Recognition 34: 1491–1501.

9. Rehm M, Bee N, André E (2008) Wave like an Egyptian: accelerometer based gesture recognition for culture specific interactions. In: HCI 2008 Culture, Creativity, Interaction. pp. 13–22.

10. Howard N, Pollock R, Prinold J, Sinha J, Newham D, et al. (2013) Effect of impairment on upper limb performance in an ageing sample population. In: Stephanidis C, Antona M, editors, Universal Access in Human-Computer Interaction. User and Context Diversity, Springer Berlin Heidelberg, volume 8010 of *Lecture Notes in Computer Science*. pp. 78-87. doi:10.1007/978-3-642-39191-0_9. URL `http://dx.doi.org/10.1007/978-3-642-39191-0_9`.

11. Arieli Y, Freedman B, Machline M, Shpunt A (2010) Depth Mapping Using Projected Patterns. US Patent: US20100118123 A1.

12. Microsoft (2011). Kinect for windows SDK beta. URL `http://www.microsoft.com/en-us/kinectforwindows/`.

13. Microsoft (2012). MSDN Library: Coordinate spaces.

14. Almeida QJ, Wishart LR, Lee TD (2002) Bimanual coordination deficits with Parkinson's disease: the influence of movement speed and external cueing. Movement Disorders 17: 30–37.

15. Helsen WF, Elliott D, Starkes JL, Ricker KL (2000) Coupling of eye, finger, elbow, and shoulder movements during manual aiming. Journal of Motor Behavior 32: 241–248.

16. Elgendi M, Picon F, Magenant-Thalmann N (2012) Real-Time speed detection of hand gesture using Kinect. In: Proc. Workshop on Autonomous Social Robots and Virtual Humans, The 25th Annual Conference on Computer Animation and Social Agents (CASA 2012), Singapore.

17. Baak A, Müller M, Bharaj G, Seidel HP, Theobalt C (2013) A data-driven approach for real-time full body pose reconstruction from a depth camera. In: Consumer Depth Cameras for Computer Vision, Springer. pp. 71–98.

18. Gorman WP, Cooper R, Pocock P, Campbell MJ (1986) A comparison of primidone, propranolol, and placebo in essential tremor, using quantitative analysis. Journal of Neurology, Neurosurgery & Psychiatry 49: 64–68.

19. Hunter J. Matplotlib. URL `http://matplotlib.sourceforge.net/`.

Tables

Table 1. Error rates for non-filtered and filtered arm-movement signals. Classification error has been calculated within (intra) and between (inter) speed classes. In case of intraclass speed classification, the lowest error is achieved by either non-filtered or filtered of feature f_1 is 0.49%, while the lowest error achieved in case of interclass speed classification is 5.43% by filtered feature f_1. Feature filtration is done using a Butterworth low-pass filter with a cutoff frequency of 6 Hz.

		Hand				**Wrist**			
		f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
		error(%)	error(%)	error(%)	error(%)	error(%)	error(%)	error(%)	error(%)
Intra	**Non-filtered**	**0.49**	0.98	58.27	24.69	1.48	4.44	55.80	31.85
	Filtered	**0.49**	2.46	58.27	11.35	1.48	4.44	55.80	15.55
Inter	**Non-filtered**	8.39	6.41	60.74	30.86	10.37	12.34	58.02	41.97
	Filtered	8.39	**5.43**	60.74	22.71	10.86	9.38	58.02	26.17
		Elbow				**Shoulder**			
		f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
		error(%)	error(%)	error(%)	error(%)	error(%)	error(%)	error(%)	error(%)
Intra	**Non-filtered**	1.48	6.41	53.82	21.72	4.44	11.60	36.04	12.83
	Filtered	1.48	5.18	53.82	11.85	4.44	11.85	36.04	15.30
Inter	**Non-filtered**	11.35	17.53	51.60	34.32	27.40	31.35	38.51	33.58
	Filtered	11.35	14.56	51.60	27.90	27.40	31.11	38.51	32.09

Table 2. Leave-one-out (LOO) crossvalidation results for the best feature for intraclass and interclass for the hand joint speed analysis. The best feature for intra- and inter- class anaysis is selected based on the overall joints analysis shown in Table 1. Each LOO step means one subject is exculded for testing and the training is done over the rest 26 subjects. The result of testing is displayed in the Error column.

	Intraclass			**Interclass**		
	Non-Filtered f_1			**Filtered** f_2		
LOO step	$\mathbf{THR_1}$	$\mathbf{THR_2}$	**Error (%)**	$\mathbf{THR_1}$	$\mathbf{THR_2}$	**Error (%)**
1	0.57	1.47	53.33	0.43	1.25	6.66
2	0.57	1.48	60.00	0.43	1.25	0.00
3	0.56	1.48	53.33	0.43	1.25	6.66
4	0.58	1.47	26.66	0.43	1.25	0.00
5	0.59	1.52	20.00	0.47	1.25	40.00
6	0.58	1.51	13.33	0.43	1.25	0.00
7	0.59	1.51	0.00	0.48	1.25	13.33
8	0.59	1.51	0.00	0.43	1.25	20.00
9	0.58	1.50	6.66	0.43	1.25	0.00
10	0.58	1.49	13.33	0.43	1.25	0.00
11	0.58	1.49	6.66	0.47	1.25	13.33
12	0.57	1.50	13.33	0.43	1.25	13.33
13	0.58	1.48	13.33	0.43	1.25	6.66
14	0.57	1.48	66.66	0.43	1.25	0.00
15	0.58	1.47	20.00	0.43	1.25	0.00
16	0.58	1.49	26.66	0.43	1.25	6.66
17	0.58	1.50	6.66	0.43	1.25	0.00
18	0.58	1.50	13.33	0.43	1.25	6.66
19	0.59	1.50	0.00	0.43	1.25	6.66
20	0.58	1.50	0.00	0.43	1.25	0.00
21	0.58	1.49	6.66	0.43	1.25	0.00
22	0.58	1.50	0.00	0.43	1.25	0.00
23	0.58	1.49	0.00	0.43	1.25	6.66
24	0.57	1.50	33.33	0.43	1.25	6.66
25	0.59	1.50	0.00	0.43	1.25	0.00
26	0.58	1.48	20.00	0.43	1.25	0.00
27	0.57	1.49	53.33	0.43	1.25	6.66

Figures

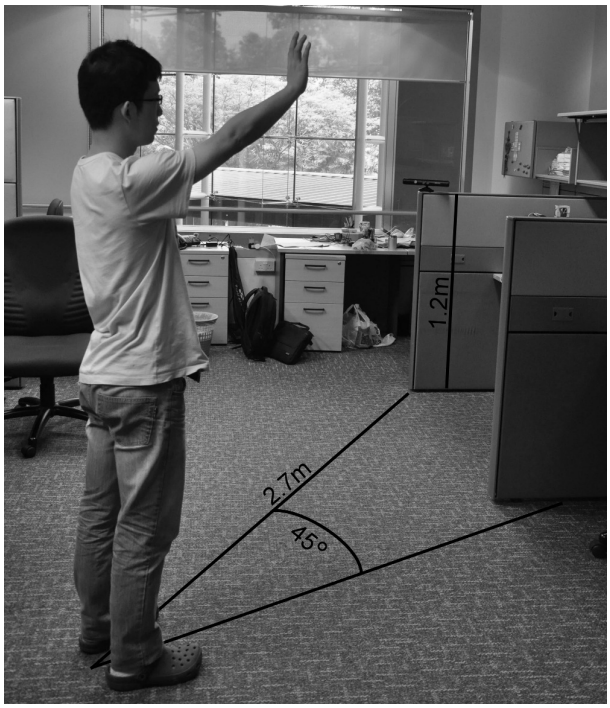

Figure 1. Experimental Setup: the user is facing the camera at angle of 45° to the right of the sensor. Every arm movement is recorded at a fixed 2.7 m distance from the camera; where the camera is placed at height of 1.2 m above the floor.

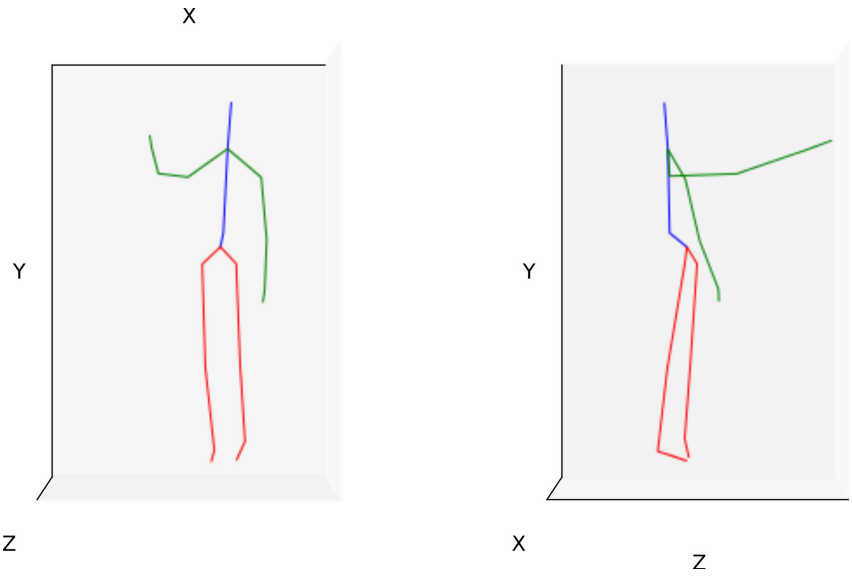

Figure 2. Front and lateral view, of a subject, computed from the sensor data. This plot represents the middle of the motion and was traced using Python 2.7 and the plotting module Matplotlib 1.1.0 [19]. The instantaneous velocity will be calculated using the x, y, z coordinates shown in the figure. The green lines represent arms, red represents legs and blue represents the torso.

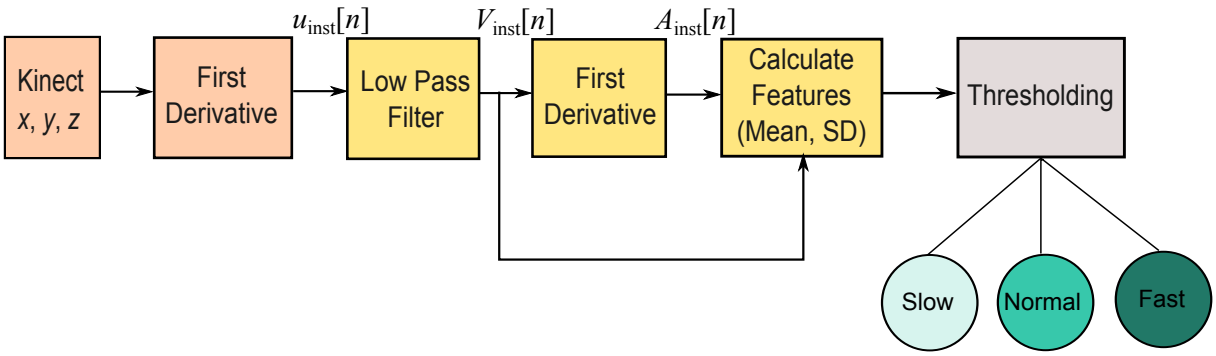

Figure 3. Flowchart for the arm-movement type classification. This is the proposed algorithm that consists of three main stages: pre-processing (importing Kinect signals and first derivative), feature extraction (lowpass filter, first derivative, and calculating features) and classification (thresholding).

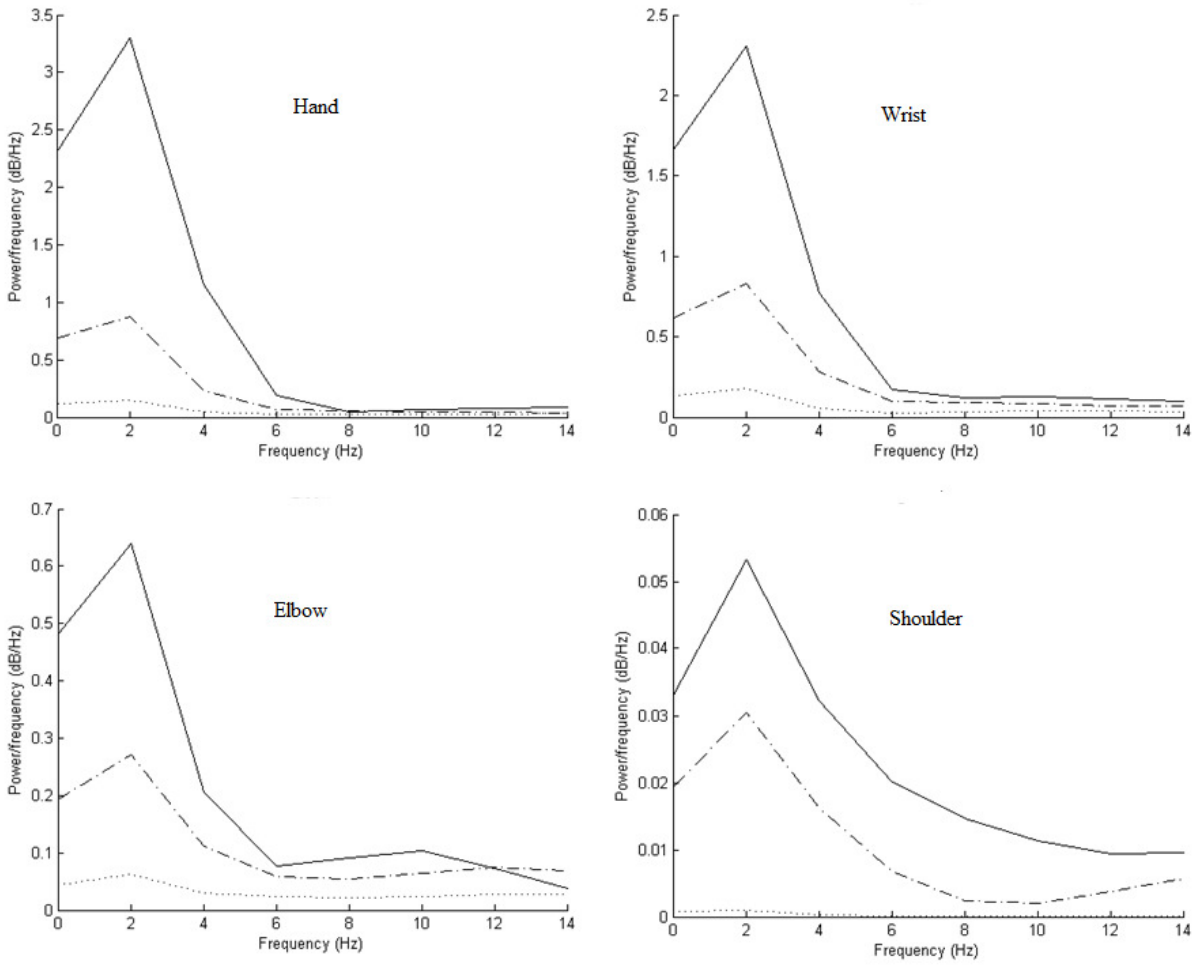

Figure 4. Power spectra of the three speed motions: slow, normal and fast. The dotted cureve represents the PSD of a slow hand movement, while the dotted curve cureve represents the PSD of a medium hand movemnt. The PSD of a fast hand movement is the solid curve.

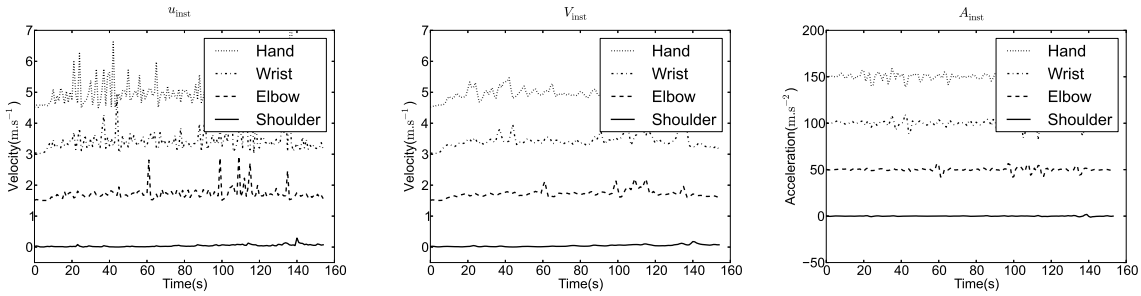

Figure 5. Comparison of instantaneous velocity and acceleration, non filtered (at left) and filtered (at right) of a slow motion for four joints: shoulder, elbow, wrist and hand of the right arm. The plots are done for one motion of one subject. For a clearer graph, extra vertical space has been added between the plots, however the scale ratio has been preserved. From bottom to top are shoulder, elbow, wrist and hand respectively. The cutoff frequency of the Butterworth low-pass filter is 6 Hz.

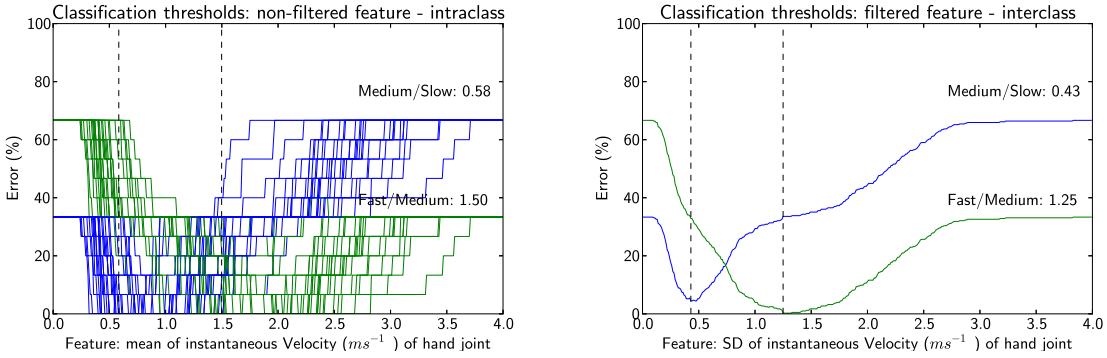

Figure 6. Threshold of classification. The figures represents, left to right, the intraclass speed classification for mean (f_1) and the interclass speed classification for SD (f_2) of the hand instantaneous velocity. The intraclass figure is made by superimposing the figure from each subject and computing thresholds using the average values of each subjects thresholds. The two dashed lines point to the two valleys in figure; and their x-axis values are the used thresholds.

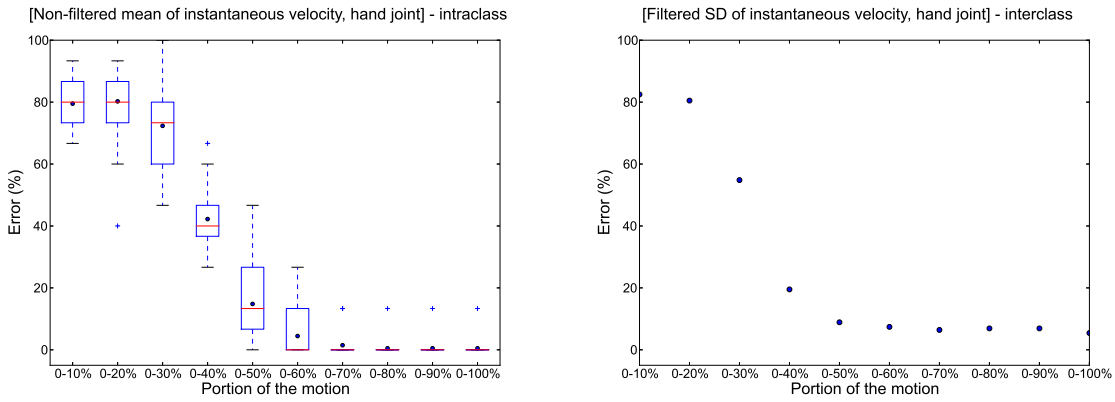

Figure 7. Classification error rate of the speed types based on the percentage of the whole motion used. This analysis was carried out to search for the most useful perentage of the whole motion that distinguishes the speed types. Knowing this percentage can lead to an effective prediction. The two figures represent, from left to right, the intraclass speed classification for mean (f_1) and the interclass speed classification for the SD (f_2) of instantaneous velocity. The intraclass boxplot shows the variation within subjects. The interclass scatter shows the exact error rate over all subjects. The portion 0–50% presents a comparable error rate to the whole motion.

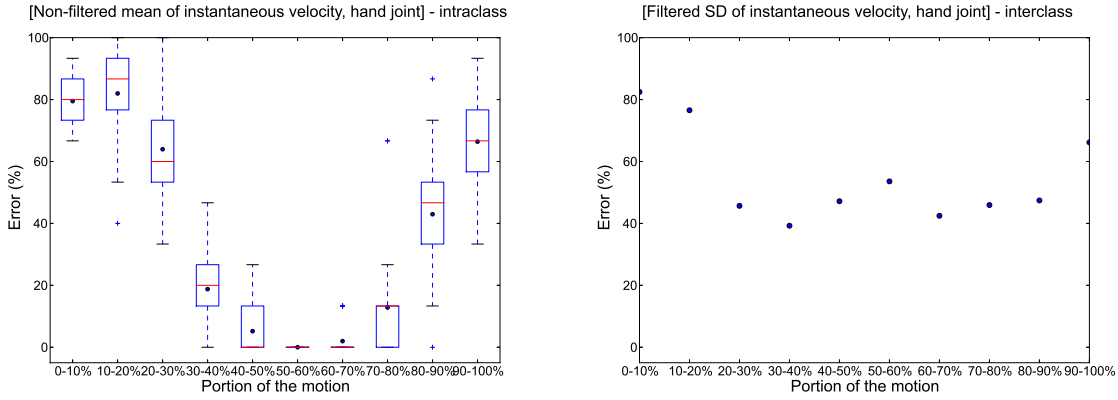

Figure 8. Classification error rate of the speed types based on a sequential 10% cuts of the whole motion. This analysis was carried out to search for the most useful 10% portion of the motion's signal that distinguishes the speed types. Knowing this percentage can lead to an effective prediction. The two figures represent, from left to right, the intraclass speed classification for mean (f_1) and the interclass speed classification for SD (f_2) of the hand instantaneous velocity. The intraclass boxplot shows the variation within subjects, while the interclass scatter shows the exact error rate over all subjects. First we can observe that the classification error diminishes in the middle of the curve, this seems to indicate that the most meaningful section of the motion is at the middle. The smallest error in the portion 50–60% for both intra- and inter- class speed classification.

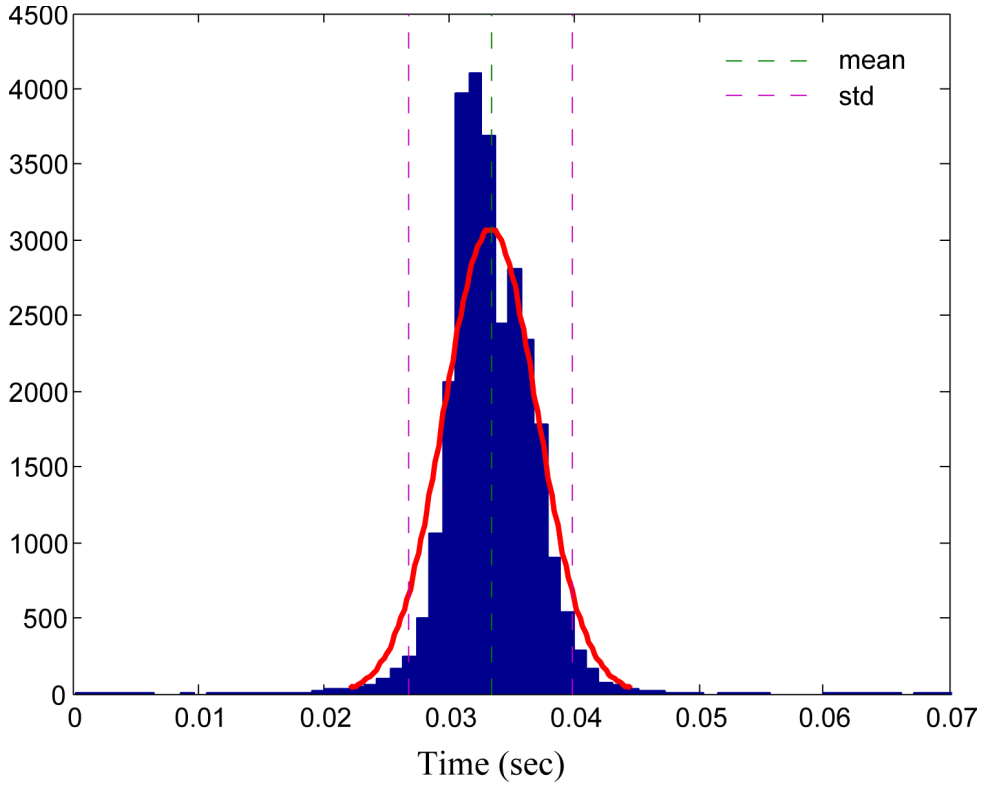

Figure 9. Histogram of timestamps difference for motion samples collected by Kinect camera. The histogram shows the statistical analysis of the time difference for motion samples collected by Kinect. The mean sampling time is 0.0334, which is 29.94 Hz, while the sampling error is 0.0037, which means sampling frequency error of 2.7 Hz. A Gaussian curve (red color) has been fitted on the data.