# Non-Invasive Motion Analysis for Stroke Rehabilitation using off the Shelf 3D Sensors

Nadav Eichler
Department of Computer Science
Haifa University, Israel
Email: eichler@outlook.com

Hagit Hel-Or
Department of Computer Science
Haifa University, Israel
Email: hagit@cs.haifa.ac.il

Ilan Shmishoni
Department of Information Systems
Haifa University, Israel
Email: ishimshoni@is.haifa.ac.il

Dorit Itah
Occupational Therapy Unit
Galilee Medical Center
Naharia, Israel
Email: dorit.itah@gmail.com

Bella Gross
Galilee Medical Center
Azrieli school of Medicine
Bar Ilan University, Israel
Email: bgross@013net.net

Shmuel Raz
Department of Information Systems
Haifa University, Israel
Email: razshmu@gmail.com

***Abstract*—Stroke is one of the most common adult injuries, with 6.5 million stroke survivors in the US alone. We use a novel motion capture system together with machine learning tools to evaluate the standard stroke rehabilitation scale, the Fugl-Meyer Assessment (FMA). FMA involves the patient performing specific motor actions. A medical professional rates the performance and provides an FMA score. We have developed a multi depth-camera system using off the shelf consumer depth cameras. Its novelty is in its ability to perform synchronization, data integration and most importantly, calibration on the fly automatically without the need of a professional operator. The camera system tracks the subject's body and outputs a stream of skeleton representations, which allows to evaluate the patients's motor performance. Using a multi camera system rather than a single camera allows capturing motion on all sides of the patient body, as required by the FMA. The system was evaluated in a pilot study at a major hospital. Applying machine learning techniques on the skeleton streams, the system was able to correctly asses FMA scores on 2 of the standard motions with close to 100% success rate. This serves as a proof of concept for the feasibility of creating a full FMA home based assessment tool.**

## I. Introduction

Stroke is a serious medical condition which occurs when the blood flow to an area in the brain is cut off. If a stroke is not detected early enough, permanent brain damage or death may occur. Around 800,000 people a year in the US incur a stroke and there are 6.5 million stroke survivors in the US. Stroke accounts for 1 of every 20 deaths in the US and amounts to nearly 133,000 people a year [1]. Rehabilitation of stroke patients is a long and slow process. The level of rehabilitation of a patient is typically evaluated using the Fugl-Meyer Assessment (FMA) [2]. This test involves the patient performing specific motor actions. A physician or skilled medical professional rates the performance on the FMA scale and a score is derived. Thus, this score is subjective and lacks a high degree of objectivity, impartiality and sensitivity. The Fugl-Meyer Assessment (FMA) is a stroke-specific, performance-based impairment index. It is designed to assess motor functioning, balance, sensation and joint functioning in patients with post-stroke hemiplegia, (a weakness of one entire side of the body). FMA is applied clinically and is used to determine disease severity, describe motor recovery, and to plan and assess treatment.

Since stroke is a debilitating disorder, patients often find it difficult to travel to stroke rehabilitation centers for testing and thus do not receive optimal medical care. Recent advances in information and communication technologies connect specialists that are centered in urban areas with population in suburban and rural areas and thus benefit these patients. This technology, which enables treatment of patients in remote areas and in nursing homes is termed *Tele-Medicine* stemming from the use of telecommunication in order to provide health care. Under the idea of Tele-Medicine, we propose an automated system for tracking and evaluating stroke rehabilitation using a cost-effective home-based system. The system must be non-invasive, easy to use, inexpensive and can be activated in a home setting. In order to be consistent with the physician's view of the examinee and also to prevent tracking failures, at least two cameras are necessary for analyzing both sides of the patient independently. Thus, our multi-camera system based on 3D cameras is appropriate for this task. In the current study, Microsoft Kinect V2 was used. This device, released in 2013, uses IR time-of-flight technology and incorporates human tracking and skeleton representation [3], [4] (see Figure 1).

Since its release, the Kinect Camera V2 has been extensively studied for its noise statistics [5], [6], [7], tracking capabilities [8], [9], and compared with state-of-the-art and commercial human motion camera systems [8], [10], [11], [12]. The Kinect has been used in various applications such as medical applications including Parkinson Tracking [8], [13], balance disorders [14], rehabilitation [15], elderly monitoring [9] as well as sport and dance tracking and analysis [16], [17], and various other computer vision applications (see [18] for a review).

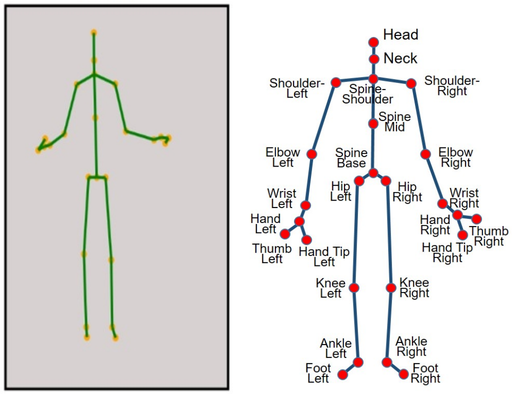


Fig. 1. Micorosft Kinect Skeleton (left) and Joint Map (right).

Several studies have considered a multi-camera setup using Kinect cameras [17], [19], [14]. The sensors either showed interference (Kinect V1) or required special spatial layout [20], or assisted calibration [14].

Several high-end motion capture systems are available such as VICON [21], Optotrak[22], Ipi Soft[23] to name a few, which provide the state-of-the-art in human body tracking performance. However, these advanced technologies have several significant limitations:

- Very expensive
- Non-Portable
- Invasive, typically require on-body markers.
- Complex to use, often requiring a professional trained operator.
- Requires system calibration typically in the form of a recording session using special dedicated accessories.
- In order to obtain a "skeleton" representation, requires additional manual measurements of subject's body parts as part of the calibration.
- Often requires a large setup area, thus is inappropriate for home settings.

Only a Few studies attempted to implement automated FMA ("Fugl-Meyer") tracking systems. In [24], the feasibility of automating FMA was tested, but its set up required a large space and was expensive (e.g., robotic arms, and EMG sensors), making it unsuitable for clinics and home setups. In another study [25], accelerometers were used for automated assessment, but it is limited in its ability to quantify many of the FMA tests due to the limited data acquired by the accelerometers [25]. Computerizing the FMA using depth sensors has previously been explored using a single Kinect v2 sensor [26] and Kinect v1 sensor [24]. In our system we use a multi-Kinect setup, which we show to be more reliable for FMA since this test requires evaluation of the patient body movement from both sides and a single Kinect is limited in the sense that occlusion and non-frontality significantly reduces body tracking capability and skeleton formation.

## II. Methods and Procedures

In a coupled study, a human motion capture system was developed [27]. This system, that will be briefly presented here, is inexpensive, portable, markerless (non-invasive), requires no calibration equipment and performs at accuracy rates on par with the competitive motion tracking systems. The developed system is targeted for easy home use and telemedicine applications, as well as serving other applications including security, object tracking, and more. The system performs run time merging of data, resulting in a more reliable and stable skeleton representation and as an added bonus, allows alignment and merging of the 3D point clouds to form a full 3D body representation.

### *A. The 3D multi-camera system*

The multi-camera tracking system, is non-invasive (no markers), inexpensive, portable and easy to use. It outputs a reliable skeleton and a merged point cloud. As in any multi-camera system, several necessary challenges must be dealt with:

- Temporal Synchronization - ensures frames from different cameras are temporally aligned prior to merging their data. For our system, we developed a unique solution for handling several computers connected to multiple Kinects simultaneously. The temporal synchronization between recordings is obtained using an NTP server. All data required for calibration and merging is transmitted to the server that runs the system algorithms. Since this requires only the skeleton data stream, which is of narrow bandwidth, the system runs in real-time over a LAN communication network producing a single fused skeleton stream (Figure 2).

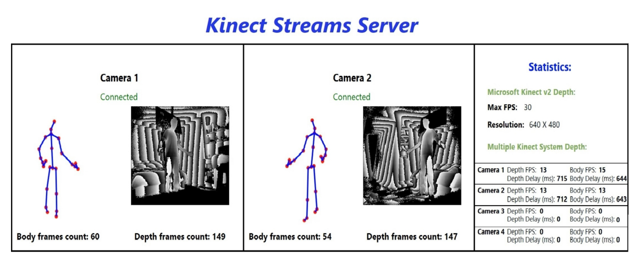


Fig. 2. Two Kinects during real-time network streaming.

- Inter-camera calibration (camera pose estimation) - multiple camera setups require scene to camera calibration as well as pose estimation between cameras in order to merge their synchronized frames. Intra-camera calibration involves mapping scene coordinates to camera coordinates. In our system, this is inherently given by the 3D cameras. However, the pose estimation or inter-camera calibration must still be calculated. There are several techniques for calibrating multiple cameras using accessories such as a 3D board [23], a checkerboard [28], a flashlight [23], and markers [21] (Figure 3). However, running calibration sessions, and using special calibration

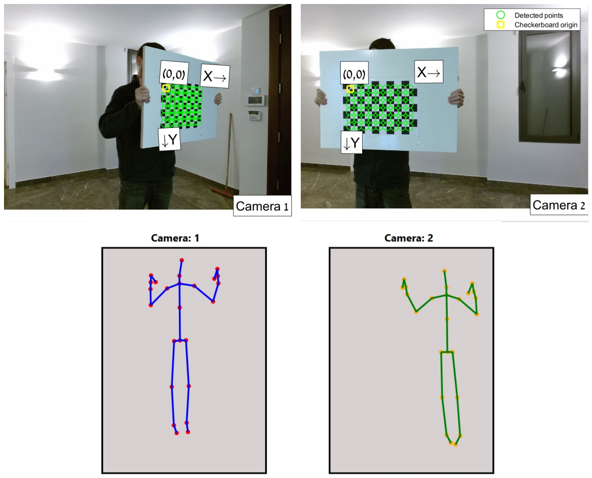


Fig. 3. Camera calibration. Using a checker board (top) and using our novel skeleton calibration (bottom).

equipment is inappropriate in our case where simplicity of activation is necessary especially for home use of the system. We propose a novel method for calibration by robustly matching skeleton joints which are generated by the camera software [3]. It can be run on-the-fly and requires no special calibration session. Rather, it relies on the body of the subject being captured in the scene.

- Data Merging - Following the calibration and the pose estimation, all the acquired data from all cameras must be aligned in a common coordinate frame. The data must then be merged into a single coherent representation. In our system, this is efficiently and reliably performed with both the body-skeleton stream and the 3D cloud points (see figures 4 and 5).

Additional details on the system can be found in [27].

### *B. Medical experiment*

We implement the multi-camera motion capture system in a medical setting for automating the FMA assessment procedure [2] in a home-setting. Our FMA application uses a two camera setup (Figure 6), ensuring each side of the patient's body is properly viewed, and producing a reliable data to be used in analysis. We conducted an Helsinki-approved study in a major public hospital using our multi-camera tracking system. 22 participants were filmed during their Fugl-Meyer assessments. The participants were twelve stroke patients and ten control healthy subjects. The subjects performed the Fugl-Meyer assessment in the hospital testing room under the guidance of a medical professional, one of the authors, who also provided the FMA rating and FMA score for the patients and the healthy subjects. The subject performed the FMA motions with the hand on the affected side as well as the hand on the unaffected side (termed- healthy hand). Each motion was repeated several times.

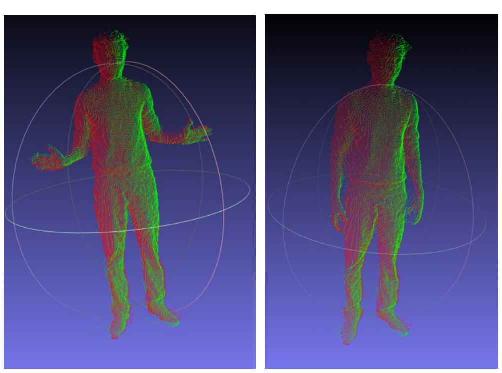

Fig. 4. Aligned and merged cloud points from 2-Cameras.

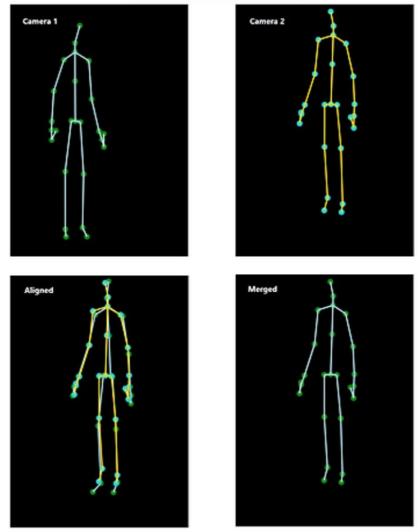


Fig. 5. Aligned and merged body skeletons from 2-Cameras. Top: camera outputs. Bottom: aligned skeleton (left) and merged (right).

The two Kinect cameras were set up in the testing room so that they did not interfere with the testing yet obtained unobstructed views of the subject. The cameras were positioned at 45 degrees angle to the subject's front view, at a distance of approximately 1.5-2 meters (see Figure 6). The two cameras recorded the body skeleton of the subject performing the motor task. In our experiment we focused on two key motor movements from the Fugl-Meyer scale protocol, the "Salute" and "Hand lift" (90 degrees) tests (see Figure 7).

An analysis application was developed for extracting mea-

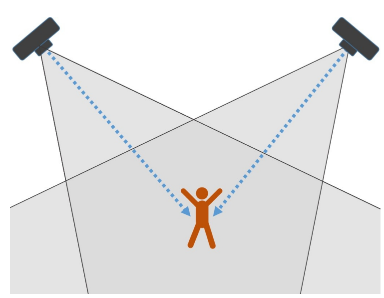

Fig. 6. The 2-camera Kinect setup [23]

Fig. 7. Fugl-Meyer Salute and Hand Lift tests.

| Measurements (participating joints) |
|---|
| Angle (Shoulder - Elbow - Wrist) |
| Distance (Head - Hand) |
| Angle (Spine Shoulder - Shoulder - Elbow) |
| Height (Hand) |
| Height (Elbow) |
| Height (Wrist) |
| Angle (Shoulder - Spine Shoulder - Spine Mid) |
| Angle (Hip - Shoulder - Elbow) |
| Distance (Hand - Knee) |
| Distance (Wrist - Hip) |
| Angle (Knee - Hip - Spine Mid) |
| Distance (Head - Elbow) |
| Distance (Camera to Shoulder) |

TABLE I
LIST OF MEASUREMENTS USED IN OUR TESTING

surements from the tracked body skeleton recordings as shown in Figure 8, and these measurements were then used to detect correlations with the physician's diagnosis. Measurements were extracted from each frame of the acquired skeleton sequence, for each of the two FMA movements. Measurements were derived from the angle defined by three skeleton joints, the distance between a pair of joints or the height moved by a single joint. The list of measurements is given in Table I. From these measurements per frame, a list of features was calculated for the whole skeleton sequence thus, providing a feature vector for each sequence, per each motion and per each subject. The extracted features per sequence included:

- Sequence time length
- Minimum, maximum of each measure within the sequence.
- Average and variance of each measure within the sequence.
- Difference between start and end values of each measure along the sequence.
- Average speed and acceleration of each measure as it changes along the sequence.

The feature vectors associated with each skeleton sequence used in the analysis and the resulting analysis are described in the following section.

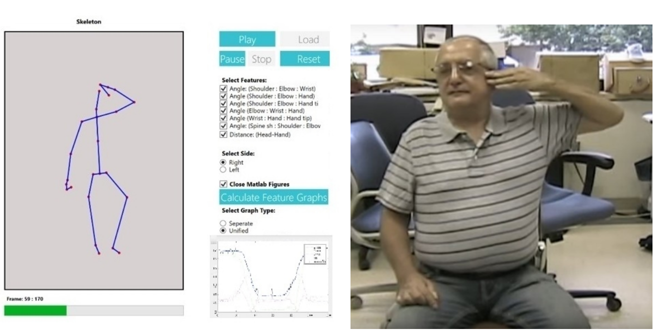


Fig. 8. Analysis of the Fugl-Meyer "Hand Salute" test.

## III. Results

In this section we analyze the data collected during the medical trial. We show that our results concur with the medical guidelines defined by the medical specialists, including the known significant movement features, the differences between patients and healthy subjects behavior and the special case of motor compensation in stroke patients. We used machine learning tools to perform classification and feature ranking. The results are presented in the following three sub sections and include: classifier results, showing success rate at predicting FMA scores, feature selection, presenting the most significant features found by the classifiers, and statistical analysis for detecting motor compensation in the stroke patients movement.

### A. Classifiers benchmark

Participants in the experiment performed several repetitions of the "Hand Lift" and "Salute" motions using each of the hands. Stroke patients typically suffer from weakness of one entire side of the body (Hemiplegia), but also suffer from a general cognitive deterioration and reduced motor ability that influences all body functioning including the "healthy" side of the body. Thus, in order to distinguish between healthy subjects and patients, we analyze each side of the subject separately as well as analyzing the asymmetry of the subject's motor performance on both sides. Towards this goal we performed several classification tests on the collected data:

1) "Raw data" - each repeated movement on each body side is collected separately as a single sample for classification. Thus, each subject has numerous samples for classification for each movement.
2) "Repetition averaging" - repetitions are averaged on each body side thus, each subject has two samples for classification for each movement, one for each body side.
3) "Asymmetry measure" - repetitions are averaged on each body side and then a measure of asymmetry is applied: $\frac{\|Right-Left\|}{\|Right+Left\|}$. Thus, each subject has one sample per movement for classification.

To test the effect of using our multi camera setup compared to a single camera setup, we analyzed these three data sets under 4 possible camera setups and data merging methods:

(a) Single camera - one of the two cameras in the setup was randomly chosen for each subject as the source camera for data collecting and only data from this camera was used in the analysis.
(b) Two cameras - following the merging of skeletons as described in [27], we consider 3 possible analyses:
  (i) Averaging - Where the data from both cameras are averaged.
  (ii) Best choice - Where the best camera data is selected. In this case the data is that which is acquired by the closest camera to the skeleton joint (subject's hand).
  (iii) All samples - Where the data from both cameras are not merged and used as sperate samples in the classification.

In order to classify between patients and healthy subjects, we ran several classifications based on the 3 types of data and the 4 types of camera acquisitions. The features used per sample were those detailed in Table I, and the class label of each sample was the FMA score provided by the physician (in our case score of 0-1 vs score of 2-3).

To build the classifiers, we used SVM [29], Single Decision Tree (C4.5) [30] and the Random Forest [31] which uses a forest of decision trees and often achieves better results. In preliminary tests, while analysing several setups and data types, SVM did not achieve better classification results than random forests. For example for the "all samples" setup for the "raw" data type for both salute and hand lift movements, SVM achieved 87.63% success rate while Random Forest achieved 90.9%. Another goal we have in the classification process is to validate the medical guidelines that were considered in feature extraction so their final ranking in classifier is very interesting for us. Decision trees are more convenient than SVM in that matter because of its built-in feature selection and that is why we focus on these algorithms in our classification analysis. Furthermore, given the ease of feature ranking using Random Forest, we report our results using only the decision tree classification approach.

| **Camera setup** | **Data used** | **Samples** | **C4.5** | **R-F** |
|---|---|---|---|---|
| All samples | Raw data | 214 | 78.50 | 82.71 |
| Best choice | Raw data | 134 | 69.40 | 82.08 |
| Single camera | Raw data | 130 | 64.61 | 77.69 |
| Averaging | Repetitions average | 44 | 77.27 | 79.54 |
| Best choice | Repetitions average | 44 | 88.63 | 79.54 |
| Single camera | Repetitions average | 45 | 66.66 | 82.22 |
| Averaging | Asymmetry measure | 22 | 36.36 | 77.27 |
| Best choice | Asymmetry measure | 22 | 81.81 | 90.90 |

TABLE II
CLASSIFIERS BENCHMARK SUMMARY - HAND LIFT

| **Camera setup** | **Data used** | **Samples** | **C4.5** | **R-F** |
|---|---|---|---|---|
| All samples | Raw data | 149 | 90.60 | 97.31 |
| Best choice | Raw data | 95 | 88.42 | 96.84 |
| Single camera | Raw data | 87 | 82.75 | 98.85 |
| Averaging | Repetitions average | 44 | 93.18 | 97.72 |
| Best choice | Repetitions average | 44 | 97.72 | 100 |
| Single camera | Repetitions average | 43 | 97.67 | 100 |
| Averaging | Asymmetry measure | 22 | 77.27 | 63.63 |
| Best choice | Asymmetry measure | 22 | 45.45 | 59.09 |

TABLE III
CLASSIFIERS BENCHMARK SUMMARY - SALUTE

| **Camera setup** | **Data used** | **Samples** | **C4.5** | **R-F** |
|---|---|---|---|---|
| All samples | Raw data | 363 | 77.95 | 90.90 |
| Best choice | Raw data | 229 | 84.71 | 89.95 |
| Single camera | Raw data | 217 | 77.88 | 87.09 |
| Averaging | Repetitions average | 88 | 85.22 | 90.90 |
| Best choice | Repetitions average | 88 | 82.95 | 93.18 |
| Single camera | Repetitions average | 88 | 90.90 | 81.81 |
| Averaging | Asymmetry measure | 44 | 70.45 | 70.45 |
| Best choice | Asymmetry measure | 44 | 72.72 | 75.00 |

TABLE IV
CLASSIFIERS BENCHMARK SUMMARY - BOTH SALUTE & HAND LIFT

The decision tree was built using three folds, with batch size 100, and confidence factor 0.25. The Random Forest classifier used the same parameters and up to 100 trees. A leave-one-out strategy was used for cross-validation, which entailed leaving out one subject at each iteration and training on the rest

Tables II and III show classification results for the Hand Lift and for the Salute motion respectively. Table IV shows classification results when data of both motions were used collectively. The tables present the summarized benchmarks across all subjects comparing the different data formats and different camera setups.

Analyzing the results shown in the tables, it can be seen that the Salute movement shows better classification results than the Hand Lift movement, with correct percentage of 96-100% (Table III rows 1-6). The success rate for the Hand Lift motion (Table II) is much lower. This may be explained by the difficulty and complexity of the Salute movement for patients. When using all data from both movements (Table IV) good performance is also achieved at 87-93% success rate.

The results also show that classifying using the asymmetry measure yields inferior results compared to the other data sets. This, and the fact that the classifications were run on data from both the affected and healthy body sides, imply that both the stroke affected side and the subject's healthy side provide distinguishing features enabling separation of patients from healthy subjects. In the context of the cameras setup, it can be seen that the 2-camera setup outperforms the single camera setup in most cases. The average correct percentage of all cases with a single-camera is 87.5% compared to 90.4% in the two-camera setup (excluding the asymmetry versions). Advantage is seen when using the best choice camera over the averaging method.

### B. Feature selection

Classification between healthy subjects and stroke patients was shown to performed better, when analyzing each side of the subject independently rather than using the asymmetry measure. Thus, in this section we focus on these high-performance classifiers and explore them further by ranking their features.

We use two feature selection ranking methods:

1) The Information Gain univariate feature ranking (Kullback-Leibler divergence) [32].
2) Random Forest features ranking [31] which summarizes the ranks of all selected features according to their positions in the decision trees.

Tables V and VI show rankings of features from the 6 best classifiers (Tables IV, II, III) using the info gain ranking and the random forest ranking respectively. The top 10 features are listed. It can be seen that the 3 most significant features across almost every classifier are (marked by colored rows):

1) Total time - Representing the period of time between start and end of the movement.
   Patients are typically slower than healthy subjects so the significance of this feature is clear.
2) Distance (Head - Elbow) Average speed - Representing the speed of the "changing distance" between the head and the elbow while performing the hand lift or salute movement. These two joints are central to performing both movements and the speed in which the movement is performed is reflected in the change in distance between these joints.
3) Angle (Hip - Shoulder - Elbow) Average speed - This feature defines the speed of the changes in the "shoulder" angle (arm to body angle) which is intuitively the significant angle while estimating the hand lifting or salute movements.

All these top features represent, in some form, the difference of "speed" between healthy and patients. The specific skeleton joints represented in these features are directly related to the medical guidelines of the FMA and the instructions given by the specialist to the patient (e.g. for the Hand lift, the patient is required to lift hand to 90° angle between arm and body).

We summarize the experimentation in concluding that FMA can be automatically scored using the features above, to distinguish between patients with high severity and low severity (and healthy) FMA scores.

### C. Compensation Statistical analysis

Motor Compensation refers to the alternative strategies developed by stroke patients in order to compensate for their difficulty or inability to perform a motor task [33]. In the context of FMA, this is expressed as increased movement signals in body parts that are unrelated to the motor task, such as the movement of the spine or the shoulders during hand lifting [33]. In the current study, we analyzed the measured motion and position of stroke patient body parts to uncover motor compensation in patients during FMA.

Following the medical guidelines [33], we analyzed the following motion features in order to detect well known compensation strategies used by stroke patients when performing Hand Lift and Salute motion:

1) Elbow Angle - (Shoulder - Elbow - Wrist) -
   Tests if the hand is straight during the movement.

| **All Samples** | **Raw Data** | **Salute & Hand lift** |
|---|---|---|
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 0.60 |
| Distance (Head - Elbow) - Average Speed | | 0.39 |
| Distance (Head - Hand) - Variance value | | 0.32 |
| Distance (Head - Hand) - Start-Stop Difference Value | | 0.28 |
| Distance (Head - Elbow) - Variance Value | | 0.27 |
| Distance (Head - Elbow) - Start-Stop Difference Value | | 0.26 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 0.26 |
| Distance (Hand - Knee) - Max Acceleration | | 0.23 |
| Height (Hand) - Min Speed | | 0.21 |
| Distance (Hand - Knee) - Average Speed | | 0.21 |
| | | |
| **Best choice** | **Repetitions avg** | **Salute & Hand lift** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 0.68 |
| Distance (Head - Elbow) - Average Speed | | 0.48 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 0.46 |
| Distance (Hand - Knee) - Average Speed | | 0.45 |
| Angle (Hip - Shoulder - Elbow) - Min Speed | | 0.40 |
| Distance (Head - Elbow) - Max Speed | | 0.40 |
| Distance (Hand - Knee) - Max Acceleration | | 0.33 |
| Distance (Head - Elbow) - Variance value | | 0.32 |
| Angle (Shoulder-Spine Shoulder-Spine Mid) - Avg Speed | | 0.31 |
| Distance (Head - Elbow) - Start-Stop Difference value | | 0.31 |
| | | |
| **All samples** | **Raw data** | **Hand lift** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 0.42 |
| Distance (Head - Elbow) - Average Speed | | 0.29 |
| Angle (Hip - Shoulder - Elbow) - Variance value | | 0.27 |
| Distance (Head - Hand) - Variance value | | 0.24 |
| Height (Hand) - Start-Stop Difference value | | 0.22 |
| Height (Wrist) - Start-Stop Difference value | | 0.22 |
| Distance (Head - Hand) - Average Speed | | 0.21 |
| Distance (Head - Elbow) - Start-Stop Difference value | | 0.20 |
| Distance (Head - Hand) - Start-Stop Difference value | | 0.20 |
| Height (Hand) - Variance value | | 0.18 |
| | | |
| **Best choice** | **Repetitions Avg** | **Hand lift** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 0.64 |
| Angle (Spine Shoulder - Shoulder - Elbow) - Max Accel. | | 0.38 |
| Angle (Hip - Shoulder - Elbow) - Start-Stop Difference value | | 0.34 |
| Angle (Hip - Shoulder - Elbow) - Variance value | | 0.33 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 0.33 |
| Angle (Shoulder - Elbow - Wrist) - Max Acceleration | | 0.33 |
| Distance (Head - Elbow) - Average Speed | | 0.32 |
| Height (Hand) - Variance Acceleration | | 0.31 |
| Height (Hand) - Max Acceleration | | 0.31 |
| Distance (Hand - Knee) - Max Acceleration | | 0.31 |
| | | |
| **All samples** | **Raw data** | **Salute** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 0.84 |
| Distance (Head - Elbow) - Average Speed | | 0.73 |
| Distance (Head - Hand) - Average Speed | | 0.62 |
| Distance (Hand - Knee) - Average Speed | | 0.61 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 0.60 |
| Distance (Head - Elbow) - Start-Stop Difference value | | 0.55 |
| Distance (Wrist - Hip) - Average Speed | | 0.51 |
| Angle (Shoulder - Spine Shoulder - Spine Mid) - Max Accel. | | 0.44 |
| Distance (Head - Elbow) - Variance value | | 0.44 |
| Distance (Head - Hand) - Start-Stop Difference value | | 0.41 |
| | | |
| **All Samples** | **Raw Data** | **Salute** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 0.99 |
| Distance (Head - Elbow) - Average Speed | | 0.86 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 0.77 |
| Angle (Shoulder - Spine Shoulder - Spine Mid) - Average Speed | | 0.70 |
| Distance (Head - Elbow) - Start-Stop Difference value | | 0.68 |
| Angle (Hip - Shoulder - Elbow) - Min Speed | | 0.64 |
| Distance (Head - Hand) - Average Speed | | 0.64 |
| Distance (Head - Elbow) - Max Speed | | 0.58 |
| Distance (Hand - Knee) - Average Speed | | 0.56 |
| Distance (Wrist - Hip) - Average Speed | | 0.50 |

TABLE V
FEATURE SELECTION RESULTS - INFORMATION GAIN

| All Samples | Raw Data | Salute & Hand lift |
|---|---|---|
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 25.97 |
| Distance (Head - Elbow) - Variance value | | 23.37 |
| Distance (Head - Elbow) - Start-Stop Difference value | | 21.00 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 20.71 |
| Height (Hand) - Start-Stop Difference value | | 20.63 |
| Distance (Wrist - Hip) - Average Speed | | 19.43 |
| Angle (Hip - Shoulder - Elbow) - Min Speed | | 18.90 |
| Distance (Head - Elbow) - Average Speed | | 18.51 |
| Height (Wrist) - Start-Stop Difference value | | 17.63 |
| Distance (Wrist - Hip) - Max Acceleration | | 16.89 |
| | | |
| **Best choice** | **Repetitions avg** | **Salute & Hand lift** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 22.50 |
| Angle (Hip - Shoulder - Elbow) - Min Speed | | 17.75 |
| Distance (Wrist - Hip) - Max Acceleration | | 16.00 |
| Distance (Head - Elbow) - Variance value | | 15.62 |
| Distance (Head - Elbow) - Max Speed | | 14.00 |
| Height (Hand) - Variance Speed | | 13.12 |
| Distance (Head - Elbow) - Average Speed | | 12.00 |
| Angle (Shoulder - Elbow - Wrist) - Max Acceleration | | 11.75 |
| Angle (Shoulder - Spine Shoulder - Spine Mid) - Avg Speed | | 11.18 |
| Height (Wrist) - Max Acceleration | | 11.00 |
| | | |
| **All Samples** | **Raw Data** | **Hand lift** |
| **Top 10 Ranked Features** | | **Rank** |
| Distance (Head - Elbow) - Average Speed | | 26.14 |
| Total time | | 24.61 |
| Angle (Hip - Shoulder - Elbow) - Variance value | | 18.06 |
| Distance (Head - Hand) - Variance value | | 17.03 |
| Height (Wrist) - Start-Stop Difference value | | 16.23 |
| Height (Hand) - Start-Stop Difference value | | 16.04 |
| Height (Hand) - Max Acceleration | | 16.00 |
| Distance (Head - Elbow) - Variance value | | 15.42 |
| Height (Hand) - Variance value | | 14.38 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 13.93 |
| | | |
| **Best choice** | **Repetitions avg** | **Hand lift** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 17.50 |
| Angle (Hip - Shoulder - Elbow) - Max Acceleration | | 16.50 |
| Angle (Hip - Shoulder - Elbow) - Start-Stop Difference value | | 14.25 |
| Distance (Head - Elbow) - Max Acceleration | | 13.50 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 10.50 |
| Angle (Shoulder - Elbow - Wrist) - Max Acceleration | | 9.75 |
| Distance (Hand - Knee) - Max Acceleration | | 9.50 |
| Distance (Wrist - Hip) - Max Acceleration | | 8.25 |
| Height (Wrist) - Max Acceleration | | 7.25 |
| Distance (Wrist - Hip) - Average Speed | | 7.00 |
| | | |
| **All Samples** | **Raw Data** | **Salute** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 27.28 |
| Angle (Shoulder - Elbow - Wrist) - Average Speed | | 17.25 |
| Distance (Head - Hand) - Average Speed | | 16.00 |
| Distance (Head - Elbow) - Average Speed | | 15.75 |
| Height (Elbow) - Average Speed | | 15.625 |
| Distance (Head - Elbow) - Variance value | | 14.50 |
| Distance (Head - Elbow) - Start-Stop Difference value | | 14.00 |
| Height (Hand) - Variance Speed | | 12.00 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 11.87 |
| Angle (Shoulder - Spine Shoulder - Spine Mid) - Min Speed | | 11.75 |
| | | |
| **Best choice** | **Repetitions avg** | **Salute** |
| **Top 10 Ranked Features** | | **Rank** |
| Total time | | 20.00 |
| Angle (Hip - Shoulder - Elbow) - Average Speed | | 16.00 |
| Distance (Head - Hand) - Average Speed | | 15.00 |
| Distance (Head - Elbow) - Average Speed | | 15.00 |
| Distance (Head - Elbow) - Max Speed | | 13.00 |
| Angle (Hip - Shoulder - Elbow) - Min Speed | | 12.00 |
| Distance (Head - Elbow) - Variance value | | 9.75 |
| Distance (Head - Elbow) - Min value | | 9.50 |
| Distance (Head - Elbow) - Max Acceleration | | 8.00 |
| Angle (Shoulder - Spine Shoulder - Spine Mid) - Avg Speed | | 7.75 |

TABLE VI
FEATURE SELECTION RESULTS - RANDOM FOREST

2) Spine Angle - (Knee - Hip - Spine Mid) -
   Tests if the back is straight during the movement.
3) Shoulder Distance - (Camera to Shoulder) -
   Tests if the shoulders are stable during the movement.

Patients should present increased motion signals for these features when they perform motor compensations, thus, we focus on the STD values of these features along the time course of the analyzed movement. We used the T-Test [34], to evaluate each compensation feature independently. The results, given in Table VII, show that for the Hand Lift motion, patients showed increased motion in all cases, with significant results, at threshold of 0.05, in the spine angle $(t(-2.437) = 42, p = 0.019)$ and the shoulder distance $(t(-2.105) = 35.5, p = 0.042)$. For the Salute motion (Table VIII), we exclude the Elbow angle, since saluting does not require maintaining a straight arm. In this case, only the spine angle is significant $(t(-2.851) = 35.43, p = 0.007)$, whereas shoulder distance shows an insignificant opposite trend.

| | Elbow Angle | Spine Angle | Shoulder Distance |
|---|---|---|---|
| Patient | 8.80±.88 | 3.37±.31 | 0.03±.004 |
| Healthy | 7.96±.68 | 2.44±.20 | 0.02±.002 |
| Sig (2-tailed) | 0.467 | 0.019 | 0.042 |

TABLE VII
COMPENSATION LEVELS IN STROKE PATIENTS AND HEALTHY SUBJECTS FOR THE HAND LIFT MOVEMENT

| | Spine angle | Shoulder distance |
|---|---|---|
| Patient | 2.47±.29 | 0.071±.01 |
| Healthy | 1.53±.16 | 0.074±.01 |
| Sig (2-tailed) | 0.007 | 0.879 |

TABLE VIII
COMPENSATION LEVELS IN STROKE PATIENTS AND HEALTHY SUBJECTS FOR THE HAND SALUTE MOVEMENT

## IV. DISCUSSION

In this study, a novel multi-camera tracking system was applied to evaluating motor movement of stroke patients as part of our stroke rehabilitation project and with the goal of allowing home assessment for patients. We showed very high classification rates between stroke patients and healthy subjects using our Fugl-Meyer analysis application. In addition, the top-ranking features were found to strongly relate to the Fugl-Meyer instructions and indicate the significance of speed of motion in determining the FMA score. Two movements were tested, both in the category of upper-limb function ability. The results show that a complex movement such as the Salute is a much better indicator than a simple movement such as the Hand Lifting. Our experiment also showed that the asymmetry between movements in patients' two hands is not a distinguishing factor and it is advantageous to analyze each hand independently.

We also found that patients show higher levels of compensation than healthy individuals. These results show, for the first time, that compensation can be detected and tracked using a consumer camera and suggests that in the future, such systems will be able to track and quantify in depth rehabilitation processes.

Our system showed consistent results with expensive and invasive high-end motion capture systems: our analysis showed that stroke patients move slower and take longer to perform a motor task compared to healthy subjects. This corresponds with [35] where infrared light emitting diodes (IREDS) were used invasively to show this effect. Our findings (Tables VII and VIII) also show increase in trunk flexion (spine motion) in patients attempting to move their hand to the target position compared to healthy subjects. This was found in [36] by using an optical motion analysis system, where eight infrared emitting diodes (IREDs) were placed on body landmarks of the hand, arm and trunk.

The high classification rate between stroke patients and healthy subject and the consistency with high-end systems show that, with additional effort, our system is suitable for stroke rehabilitation quantification from the patients home. Additional effort is needed in developing a dedicated user interface for a system operated by unprofessional end users.

## REFERENCES


[1] B. Benjamin, Chiuve, Cushman, Das, Deo, Ferranti, Floyd, Fornage, Gillespie *et al.*, "Heart disease and stroke statistics2017," *Circulation*, 2017.

[2] A. R. Fugl-Meyer, L. Jääskö, I. Leyman, S. Olsson, and S. Steglind, "The post-stroke hemiplegic patient. a method for evaluation of physical performance." *Scandinavian journal of rehabilitation medicine*, vol. 7, no. 1, pp. 13–31, 1974.

[3] Shotton, Fitzgibbon, Cook, Sharp, Finocchio, Moore, Kipman, and Blake, "Real-time human pose recognition in parts from single depth images," in *Int. Conf. Computer Vision Patern recognition (CVPR)*, 2011, pp. 1297–1304.

[4] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," *Communications of the ACM*, vol. 56, no. 1, pp. 116–124, 2013.

[5] Mallick, Das, and Majumdar, "Characterizations of noise in kinect depth images: A review," *IEEE Sensors Journal*, vol. 14, no. 6, pp. 1731–1740, 2014.

[6] Choo, Landau, DeVore, and Beling, "Statistical analysis-based error models for the microsoft kinect depth sensor," *Sensors*, vol. 14, no. 9, pp. 17 430–17 450, 2014.

[7] Falie and B. ., "Noise characteristics of 3d time-of-flight cameras," in *International Symposium on Signals, Circuits and Systems, ISSCS*, vol. 1, 2007, pp. 1–4.

[8] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, "Accuracy of the microsoft kinect sensor for measuring movement in people with parkinson's disease," *Gait & posture*, vol. 39, no. 4, pp. 1062–1068, 2014.

[9] S. Obdrzalek, G. Kurillo, F. Ofli, R. Bajcsy, E. Seto, H. Jimison, and M. Pavel, "Accuracy and robustness of kinect pose estimation in the context of coaching of elderly population," in *Engineering in medicine and biology society (EMBC), 2012 annual international conference of the IEEE*. IEEE, 2012, pp. 1188–1193.

[10] Clark, Pua, Bryant, and Hunt, "Validity of the microsoft kinect for providing lateral trunk lean feedback during gait retraining," *Gait & posture*, vol. 38, no. 4, pp. 1064–1066, 2013.

[11] Wang, Kurillo, Ofli, and Bajcsy, "Evaluation of pose tracking accuracy in the first and second generations of microsoft kinect," in *International Conference on Healthcare Informatics (ICHI)*. IEEE, 2015, pp. 380–389.

[12] Bonnechere, Jansen, Salvia, Bouzahouene, Omelina, Moiseev, Sholukha, Cornelis, Rooze, and V. S. Jan, "Validity and reliability of the kinect within functional assessment activities: Comparison with standard stereophotogrammetry," *Gait & posture*, vol. 39, no. 1, pp. 593–598, 2014.

[13] Plotnik, Shema, Dorfman, Gazit, Brozgol, Giladi, and Hausdorff, "A motor learning-based intervention to ameliorate freezing of gait in subjects with parkinsons disease," *J. Neurol.*, vol. 261, pp. 1329–1339, 2014.

[14] Funaya, Shibata, Wada, and Yamanaka, "Accuracy assessment of kinect body tracker in instant posturography for balance disorders," in *2013 7th International Symposium on Medical Information and Communication Technology (ISMICT)*, 2013, pp. 213–217.

[15] T. T. Liu, C. T. Hsieh, R. C. Chung, and Y. S. Wang, "Physical rehabilitation assistant system based on kinect," vol. 284, pp. 1686–1690, 2013.

[16] D. Alexiadis, P. Kelly, P. Daras, N. E. O'Connor, T. Boubekeur, and M. B. Moussa, "Evaluating a dancer's performance using kinect-based skeleton tracking," in *Proceedings of the 19th ACM international conference on Multimedia*. ACM, 2011, pp. 659–662.

[17] A. Kitsikidis, K. Dimitropoulos, S. Douka, and N. Grammalidis, "Dance analysis using multiple kinect sensors," in *International Conference on Computer Vision Theory and Applications (VISAPP2014)*, 2014, pp. 789–795.

[18] Han, Shao, Xu, and Shotton, "Enhanced computer vision with microsoft kinect sensor: A review," *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1318–1334, 2013.

[19] Maimone, Bidwell, Peng, and Fuchs, "Enhanced personal autostereoscopic telepresence system using commodity depth cameras," *Computers Graphics*, vol. 36, no. 7, pp. 791 – 807, 2012.

[20] S. Asteriadis, A. Chatzitofis, D. Zarpalas, D. Alexiadis, and P. Daras, "Estimating human motion from multiple kinect sensors," in *Proceedings of the 6th International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications*. ACM, 2013, p. 3.

[21] VICON, "Motion capture system homepage," http://www.vicon.com.

[22] OPTOTRAK, "Motion capture system homepage," https://www.ndigital.com/msci/products/optotrak-certus/.

[23] IPIsoft, "Motion capture system homepage," http://www.ipisoft.com/.

[24] P. Otten, J. Kim, and S. H. Son, "A framework to automate assessment of upper-limb motor function impairment: A feasibility study," *Sensors*, vol. 15, no. 8, pp. 20 097–20 114, 2015.

[25] J. Wang, L. Yu, J. Wang, L. Guo, X. Gu, and Q. Fang, "Automated fugl-meyer assessment using svr model," in *Bioelectronics and Bioinformatics (ISBB), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 1–4.

[26] W.-S. Kim, S. Cho, D. Baek, H. Bang, and N.-J. Paik, "Upper extremity functional evaluation by fugl-meyer assessment scoring using depth-sensing camera in hemiplegic stroke patients," *PloS one*, 2016.

[27] N. Eichler, "A Novel 3D Motion Capture System using Multiple Consumer Depth Sensors," Master's thesis, Department of Computer Science, University of Haifa, Haifa, Israel, 2017.

[28] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 1. Ieee, 1999, pp. 666–673.

[29] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.

[30] J. R. Quinlan, "C4.5: Programming for machine learning," *Morgan Kauffmann*, vol. 38, 1993.

[31] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

[32] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[33] M. F. Levin, J. A. Kleim, and S. L. Wolf, "What do motor recovery and compensation mean in patients following stroke," *Neurorehabilitation and neural repair*, vol. 23, no. 4, pp. 313–319, 2009.

[34] J. F. Box, "Guinness, gosset, fisher, and small samples," *Statistical science*, vol. 2, no. 1, pp. 45–52, 1987.

[35] M. Cirstea and M. F. Levin, "Compensatory strategies for reaching in stroke," *Brain*, vol. 123, no. 5, pp. 940–953, 2000.

[36] S. M. Michaelsen, S. Jacobs, A. Roby-Brami, and M. F. Levin, "Compensation for distal impairments of grasping in adults with hemiparesis," *Experimental Brain Research*, vol. 157, no. 2, pp. 162–173, 2004.