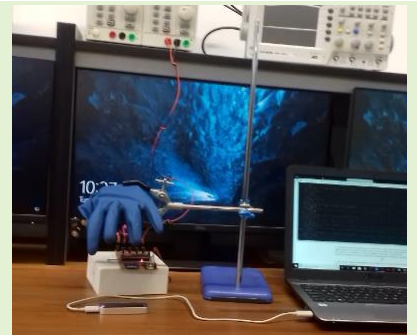


# Comparing Dynamic Hand Rehabilitation Gestures in Leap Motion Using Multi-dimensional Dynamic Time Warping

Hussein Walugembe, Member, IEEE, Chris Phillips, Member, IEEE, Jesus Requena-Carrion, and Tijana Timotijevic

**Abstract**— We propose and evaluate the use of Multi-dimensional Dynamic Time Warping (MDTW) for comparing dynamic hand rehabilitation gestures that would be performed by a patient (query) relative to hand gestures prepared by a physiotherapist (reference). MDTW enables us to determine how similar or different a query dynamic hand gesture is to a reference one whilst filtering out unwanted sources of error resulting from positional, rotational or speed differences between the query and the reference actions. It produces a minimum-distance value of a warp path after aligning a query dynamic hand gesture with a reference one. A low minimum-distance value implies the two gestures being compared are similar and high minimum-distance value implies the two gestures vary to a greater extent. When we deliberately compare a specific hand gesture with itself, we obtain a minimum-distance value of  $0^\circ$  indicating the similarity is 100%. Furthermore, when we compare two closely similar hand gestures i.e. gesture 1 and gesture 4, a minimum-distance value of  $35.9^\circ$  is obtained. However, when we compare two quite different gestures i.e. gesture 2 and gesture 3, a minimum-distance value of  $248.5^\circ$  is obtained. Therefore, a physiotherapist can establish whether a patient performs hand rehabilitation gestures satisfactorily or an adjustment is required based on the minimum-distance values of the warp paths.



**Index Terms**— Dynamic hand gesture, Hand rehabilitation, Leap Motion controller, Multi-dimensional Dynamic Time Warping.

## I. INTRODUCTION

PATIENTS suffering from stroke, neurological disorders, hand related injuries etc. usually need different forms of hand rehabilitation in order to recover expeditiously [1]. Most hand rehabilitation procedures are recommended to be executed with help of a physiotherapist that supervises a patient so that he or she can perform the correct hand exercises or gestures [2]. The cost of rehabilitation related services at hospitals or clinics and an increasing size of ageing population are responsible for considering the relocation of hand rehabilitation related services away from medical centres i.e. at home [3] to reduce expense and improve patient convenience [4]. With the aid of sensor “gadgets” and computing devices, home-based rehabilitation may be a viable option to eliminate the presence of a physiotherapist during a hand rehabilitation session [2]. However, in such a scenario, offline monitoring should be implemented to motivate patients

and make corrections to imprecise hand rehabilitation exercises [4]. This form of monitoring requires parameters to be recorded and submitted to a physiotherapist or an expert for analysis and prompt feedback [2].

Hand rehabilitative approaches are generally tailored to the patient’s needs and due to this patient-tailoring, a high resource demand is likely to be experienced. This is common to many rehabilitative treatments [5]. The resource demand may include the time a therapist spends interacting with a patient, costly instrumentation and testing equipment, the use of healthcare facilities and the related overheads [5]. In order to reduce these costs, researchers have turned their attention to implementing contactless systems with the help of computer vision techniques [2]. However, vision-based systems have limitations [6]. The Performance of such systems is hindered by factors such as the background, illumination contrast, noise related factors, ease of use etc. [6]. For instance, a scheme proposed in [6] aids rehabilitation whereby a ball is fixed to a hand gripper. Even so this may not effectively support hand finger rehabilitation since a patient is required to grasp when performing a rehabilitation procedure yet he or she may not be capable of this depending on severity of the illness [7].

There exist markerless devices in healthcare stores that can be utilised by people affected with hand injuries when executing hand rehabilitation tasks. Examples of these devices

Manuscript received June 15, 2020; revised October 7, November 18, and December 11, 2020; accepted December 20, 2020.

Hussein Walugembe is with School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, UK (e-mail: h.walugembe@qmul.ac.uk).

Chris Phillips, Jesus Requena-Carrion, and Tijana Timotijevic are also with School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, UK.

include Intel's RealSense 3-D Camera, the Inertial Measurement Unit (IMU) sensor, the Kinect Sensor [8]-[11], tilt sensors [12], [13], camera supported devices [14]-[16], Leap Motion controller (LM) [17], etc. In addition, wearable devices have also been employed in hand rehabilitation settings through implementing a number of sensors such as optical linear encoder, knitted piezoresistive fabric, smart wearable armband [18], etc. In our analysis of different sensors, we have discovered that some sensors may make patients uncomfortable when using them, at least initially, because they need time to familiarize with the setup regime before use [4]. For instance, tilt sensors typically require a patient to perform a complex start-up process [19].

During the implementation of our proposed hand rehabilitation setup, we deployed LM since it is a lightweight and affordable device compared to the most of previously mentioned sensors. Moreover, it can be easily operated by a patient as required. When connected to a computing device the LM can guide patients through hand rehabilitation exercises like the ones employed when recovering from strokes or hand related injuries, to improve the performance of activities of daily living in order to boost wellbeing [2]. In addition, a patient can use LM at any convenient location with limited investment expenditure, and without the necessary presence of a skilled person to calibrate and adjust the equipment [17].

However LM can also be applied in various areas. For instance, in an automotive industry environment [20], both LM and Kinect data were employed for programming the robot's movement and managing the program's execution in an open-ended system. In [21], LM was employed to recognize various sign languages. Furthermore, LM is widely used in applications of virtual reality (VR) that incorporate entertainment industry and education sector [22].

In this paper we compare dynamic hand rehabilitation gestures such that a physiotherapist or an expert could establish how well a patient performs hand dynamic gestures in comparison to predefined ones that are recommended by a physiotherapist. After establishing the extent to which a performed hand gesture matches a reference one, meaningful feedback can be generated to correct or encourage a patient to refine their actions during a session [23]. We propose and implement Multi-dimensional Dynamic Time Warping (MDTW) to quantify how similar or different two dynamic hand gestures are from each other [24]. Dynamic Time Warping (DTW) is a broadly applied technique in speech recognition for establishing to what extent any two time series are similar or different. Both DTW and MDTW techniques utilise a distance metric between a query time-series and a reference one and create a discriminating value: a low distance value when the two time-series are similar and a high distance value when the two time-series are different [25].

There exist other possible dissimilarity measures such as Longest Common SubSequence [24], [26], edit distance with real penalty [24], edit distance on real sequences [24], [27], and time warp edit distance [24], [28]. However, MDTW has an extra advantage since its implementation is simple and efficient. Furthermore, MDTW is superior because it is not necessary for both time-series being compared to be of equal length as required by typical distances and this behaviour is termed elasticity [24]. MDTW is therefore an elastic

dissimilarity technique that estimates the greatest match within two time-series by reducing a distance between them [24].

The novel contributions of our paper are: 1) Formulating a specific problem regarding the comparison of a dynamic query hand gesture with a dynamic reference hand gesture using DTW and MDTW; 2) Representing dynamic hand gestures using a feature vector in 3-D space; 3) Implementing and evaluating the efficacy of the MDTW technique.

The remaining sections of this paper are arranged as follows. Section II discusses potential applications and limitations of DTW. Section III presents significant content regarding the anatomy of a hand. It further presents how joint angles can be calculated. Dynamic time warping is discussed in Section IV where the problem formulation regarding both DTW and MDTW is illustrated. Section V discusses our methodology and experimental results are presented in Section VI. Finally, Section VII presents concluding remarks and future work.

## II. APPLICATIONS AND LIMITATIONS OF DTW IN LITERATURE

DTW has been implemented to measure the correlation between two brain parts by performing a comparison between the synchronization and asynchrony of the time-series [29]. DTW was employed because it considers the inherent timing and efficacy of functional magnetic resonance imaging (fMRI) time-series unlike Pearson correlation. However, in order to establish the best community structure of brain networks for each subject under consideration, other techniques were employed [29].

One of the limitations of DTW was reported by [30] when implemented for signature verification. The implementation involves obtaining correspondence and the similarity of two planar curves. The limitation was that in portions of the curves at locations with sparse sampling, insufficient resolution in the matching procedure was experienced attributed to the fact that DTW matches only individual samples instead of continuous curves. The authors in [30] further proposed that a feasible solution to this limitation is to oversample the curves. Oversampling can be implemented by employing a spline interpolation prior to curve matching.

In [31] DTW is implemented to measure time-series similarity, perform classification, and identify corresponding portions between two time-series. One strategy that was implemented applies a multilevel technique which repeatedly predicts a measurement from a coarse resolution and improves the predicted measurement. The strategy possesses linear time and space complexity and partially solves the issues of quadratic time and space complexity experienced by DTW when implemented with large time-series data sets [31].

In order to detect, locate, and characterise damage on large remote structural areas, DTW has been proposed and implemented where it compares guided wave data to a baseline signal [32]. DTW performs better compared to similar techniques such as baseline signal stretch and scale transform [32]. In fact, in this instance DTW takes into consideration factors such as large temperature differences, long propagation intervals, and high frequencies unlike baseline signal stretch and scale transform techniques. However, DTW does not completely detect structural changes through realigning guided waves to a baseline since is

sensitive to temperature changes whilst not being adequately sensitive to long propagation distances and high frequencies [32].

DTW was implemented as a technique to distinguish various mechanical diagnostics of on-load tap changer (OLTC) [33] devices. OLTC is one of the parts in a transformer that is susceptible to faults. In addition, durability of the whole power system greatly depends on the OLTC status. Experiments demonstrate that DTW can adequately diagnose various OLTC faults under diverse fault extremities [33]. On the other hand, the implemented DTW could not efficiently handle the limitations of excessive match and distance estimation [33] and this necessitates further research.

### III. BACKGROUND

We briefly provide an overview of the hand that is essential for our MDTW implementation. We further provide how we calculate finger joint angles and, lastly, define a dynamic hand gesture.

#### A. Anatomy of the Hand

We briefly introduce a hand overview here i.e. the bones of the hand and finger joints.

##### 1) Bones of the Hand

A human hand includes metacarpal bones. These are five bones that are located from the wrist to the initial bottom joint of every finger. In addition, the hand comprises fourteen small bones referred to as phalanges (phalanx for singular). Apart from the thumb that possesses only two phalanges, the other four fingers possess three phalanges i.e. distal, middle and proximal [34], [35].

##### 2) Finger Joints

Fig. 1 illustrates the joints and bones of a human hand. The finger joints arise when a pair of adjacent finger bones meet. Apart from the thumb that has two joints, the four fingers (index, middle, ring and little) have three joints. The following describes the different fingers joints.

- *Distal Inter-Phalangeal (DIP) Joint*

The DIP joint is situated at the topmost part of the finger in the vicinity of the lower part of the fingernail. This joint intersects the distal phalanx and the middle phalanx bone. Separate from the thumb, the four fingers each possess a DIP joint.

- *Proximal Inter-Phalangeal (PIP) Joint*

The PIP joint is the middle joint of each of the four fingers but not the thumb. This joint intersects the middle phalanx and the proximal phalanx bone.

- *Inter-Phalangeal (IP) Joint*

The IP joint is situated at the topmost part of the thumb in the vicinity of the base of the nail. This joint intersects the distal phalanx and the middle phalanx bone.

- *Meta-Carpo-Phalangeal (MCP) Joint*

The MCP joint is situated at the bottom part of all five fingers i.e. all the fingers possess a MCP joint. This joint intersects the metacarpal bone and the initial phalanx bone.

#### B. Joint Angle Calculation

Equation (1) and Fig. 2 provide a demonstration of how to calculate a specific joint angle of a finger of a hand. In the figure, an index finger is considered and the angle  $\theta$  is the Proximal Inter-Phalangeal (PIP) joint angle.  $u$  illustrates the direction vector that describes the middle phalanx bone and  $v$  illustrates the direction vector that describes the proximal phalanx bone.

$$\cos \theta = \frac{(\vec{u} \cdot \vec{v})}{(\|\vec{u}\| \cdot \|\vec{v}\|)} \quad (1)$$

#### C. Definition of a Dynamic Gesture

A dynamic hand gesture is represented as a sequence of postures changing over time and each posture is described by a set of joint angles. Mathematically, a dynamic hand gesture is illustrated using a set of  $K$  measured finger joint angles that evolve over time i.e. a dynamic hand gesture  $g_t$  at an instant  $t$  is described by  $g_t = \{\theta_{1t}, \theta_{2t}, \dots, \theta_{Kt}\}$ .

### IV. DYNAMIC TIME WARPING

#### A. Problem Formulation for Dynamic Time Warping (DTW)

Our objective is to establish how similar or different a query dynamic hand gesture is in comparison to a reference dynamic hand gesture whilst compensating for differences in the duration of gestures, rotation of the hand, reasonable distance from LM sensor etc. For simplicity, we initially assume the feature vector constitutes only one measured feature in order to illustrate how traditional DTW works. In later discussion, we extend DTW to Multi-dimensional Dynamic Time Warping (MDTW) so that it suits our proposed framework.

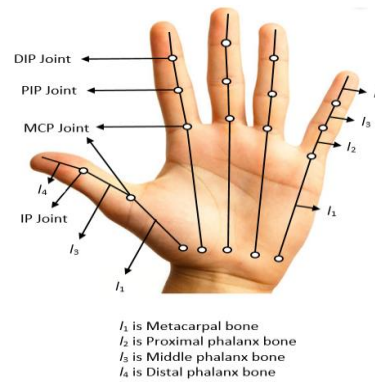


Fig. 1. Hand bones and joints.

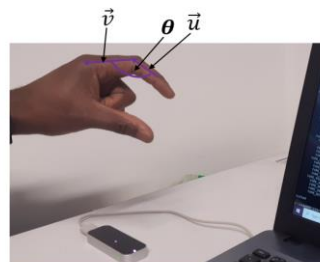


Fig. 2. Calculation of the PIP joint angle  $\theta$  of the index finger.

We let a query dynamic hand gesture and a reference dynamic hand gesture be represented by  $X$  and  $Y$ , respectively [36]. Specifically  $X = (x_1, x_2, \dots, x_M)$  where  $M \in \mathbb{N}$  and  $Y = (y_1, y_2, \dots, y_N)$  where  $N \in \mathbb{N}$ . From now onwards, we use  $i$  and  $j$  to represent an entry in time series  $X$  and  $Y$ , respectively. We define the Euclidean distance between any two samples in time series  $X$  and  $Y$  as [36],

$$d(i, j) = \sqrt{(x_i - y_j)(x_i - y_j)} \quad (2)$$

where  $1 \leq i \leq M$ , and  $1 \leq j \leq N$ .

A two-dimensional  $N$  by  $M$  cost matrix  $D$  is created and every individual value of the  $D(i, j)$  is determined as follows:

$$D(1, 1) = d(1, 1) \quad (3)$$

$$D(1, j) = D(1, j-1) + d(1, j) \quad 2 \leq j \leq N \quad (4)$$

$$D(i, 1) = D(i-1, 1) + d(i, 1) \quad 2 \leq i \leq M \quad (5)$$

$$D(i, j) = d(i, j) + \min \begin{cases} D(i, j-1) \\ D(i-1, j-1) \\ D(i-1, j) \end{cases} \quad (6)$$

Equation (6) holds for the same range of  $i$  and  $j$  as for (4) and (5) i.e.  $2 \leq i \leq M$  and  $2 \leq j \leq N$ .

Then a warping path  $W = w_1, w_2, \dots, w_r$  is an adjacent collection of some matrix constituents that are always close to the diagonal. These matrix elements when added together are equivalent to a minimum-distance of a warp path.

The minimum-distance of a warp path illustrates a mapping between  $X$  and  $Y$  that fulfils the subsequent requirements [37].

### 1) Boundary Requirements

$w_1 = (1, 1)$  and  $w_r = (M, N)$  where  $r$  is the length of the warping path.

### 2) Continuity Conditions

Given  $w_z = (a, b)$  and  $w_{z-1} = (a', b')$ , then  $a - a' \leq 1$  and  $b - b' \leq 1$  must be fulfilled.

### 3) Monotonicity Requirements

Given  $w_z = (a, b)$  and  $w_{z-1} = (a', b')$ , then  $a - a' \geq 0$  and  $b - b' \geq 0$  must be satisfied.

We can now illustrate how to obtain a minimum-distance warp path with an example. Given  $X = (1, 2, 4, 3, 5, 3, 2, 3, 2, 5)$  that represents a query series and  $Y = (1, 1, 2, 4, 3, 5, 3, 2, 3, 2)$  which represents a reference series, we can construct a cost matrix  $D$ . The minimum-distance of a warp path is traced through the matrix constituents from  $D(1, 1)$  to  $D(M, N)$ , highlighted in pink, as shown in Fig. 3. The minimum-distance of the warp path using Euclidean distance as a metric is three. The above DTW formulation can only be implemented if a framework considers a single feature

alignment i.e. one-dimensional measurements [36].

## B. Problem Formulation for Multi-dimensional Dynamic Time Warping (MDTW)

Since in our framework, we are considering up to 14 joint angles to describe a dynamic hand gesture at an instant, we employ MDTW. For MDTW, the two time-series  $X$  and  $Y$  must be initially created as multi-dimensional matrices where each row represents the time-series of a single measured feature and each column represents all the measured features at a given instant. The matrices  $X$  and  $Y$  can now be written as illustrated below.  $M$  and  $N$  are samples of dynamic query hand gesture and dynamic reference hand gesture, respectively, and  $K$  is the number of measured features being considered.

$$X = \begin{pmatrix} x_{1,1} & \dots & x_{1,M} \\ \vdots & \ddots & \vdots \\ x_{K,1} & \dots & x_{K,M} \end{pmatrix}$$

$$Y = \begin{pmatrix} y_{1,1} & \dots & y_{1,N} \\ \vdots & \ddots & \vdots \\ y_{K,1} & \dots & y_{K,N} \end{pmatrix}$$

We now define Euclidean distance between  $X$  and  $Y$  as

$$d(i, j) = \sqrt{(X_i - Y_j)^T (X_i - Y_j)}. \quad (7)$$

The entries in the cost matrix  $D$  can then be determined as shown in (4), (5) and (6). Then the MDTW algorithm searches for a minimum-distance warp path that runs close to the diagonal line from  $D(1, 1)$  to  $D(M, N)$ . The warping path must fulfil all three requirements as described in the case for DTW. If the minimum-distance of a warp path is of a lower value, then the two time-series, i.e. the two dynamic hand gestures being compared, are similar, otherwise the two dynamic hand gestures being compared are effectively different to a quantifiable extent.

## V. METHODOLOGY

In our proposed MDTW implementation, we make use of an LM device that can be readily purchased by a patient or clinic at an affordable price [38]-[40]. Fig. 4 demonstrates how to configure LM in an experimental setup. In the figure the LM device is attached to a computing device such as a laptop on the right.

2	16	8	8	4	6	2	1	1	0	3
3	15	8	6	3	5	1	1	0	1	3
2	13	7	5	3	5	1	0	1	1	4
3	12	7	3	2	2	0	1	1	2	4
5	10	6	2	2	0	2	5	5	6	4
3	6	3	1	0	2	2	3	3	4	6
4	4	2	0	1	2	3	5	6	8	9
2	1	0	2	3	6	7	7	8	8	11
1	0	1	4	6	10	12	13	15	16	20
1	0	1	4	6	10	12	13	15	16	20
	1	2	4	3	5	3	2	3	2	5

Fig. 3. Cost matrix and the minimum-distance of the warp path.

### A. Architecture of the Proposed System

We obtain an input signal when a hand is placed above the LM that is connected to a computing device using a USB cable. On the computing device, the LM application and a Java customised application are started simultaneously and frames of data that represent palm and finger movements are displayed. We then calculate joint angles of the fingers from the frames of data and, consequently, obtain feature vectors that represent various hand gestures.

After obtaining feature vectors that represent both the dynamic query hand gesture and the dynamic reference hand gesture, we apply MDTW on both feature vectors and the minimum-distance value of a warp path is obtained. The minimum-distance value signifies how similar or different the two hand gestures are. A low minimum-distance value implies the two hand gestures that are compared are similar. On the other hand, a high minimum-distance value implies the two hand gestures vary to a certain extent. Fig. 5 illustrates the architecture of the MDTW comparison process.

### B. Description of the Feature Vector

A dynamic hand gesture is represented using a set of fourteen measured finger joint angles that evolve over time. The fourteen finger joint angles that constitute the feature vector are illustrated in Table 1. Other features e.g. the angles between the finger and palm position are sufficient to be included in the feature vector. However, we choose to limit the feature vector to only finger joint angles since including other features does not contribute a considerable effect on the experimental results regarding the application of MDTW.

### C. Robotic Hand

In order to simulate the repetitive nature of dynamic hand gestures and provision of carefully controlled experimental conditions, we used a robotic hand during data collection. In addition, compared to a human hand, a robotic hand permits gestures to be performed frequently in a short time provided the movements of all the necessary servo controllers are suitably programmed for a specific hand gesture. Furthermore, the dynamic hand gestures are performed with minimum positional or temporal errors and this provides a suitable working environment to reduce “unintentional” errors during experimentation. The robotic hand was fitted with a glove so that it can be easily recognised by the LM. Fig. 6 illustrates the robotic hand performing some dynamic hand gestures.

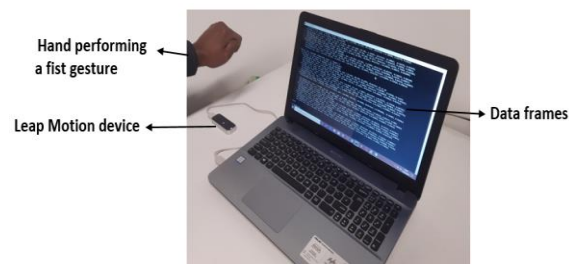


Fig. 4. A hand above the LM device in an experimental setup.

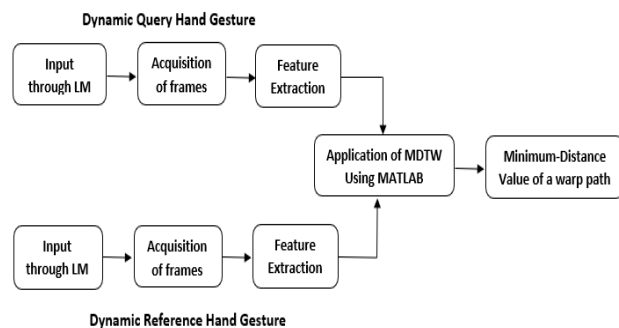


Fig. 5. Architecture of the proposed MDTW implementation setup.

### D. Gestures Under Consideration

We have considered four different dynamic hand gestures. These hand gestures are similar to those implemented in [41] for hand strengthening exercises in chronic stroke patients. Dynamic hand *gesture 1* is when a hand is performing a full fist gesture where all the five fingers move close to the palm. Dynamic hand *gesture 2* is where only the thumb and index move close to the palm whereas the rest of the 3 fingers remain stationary. Dynamic hand *gesture 3* is performed when the middle, ring and little fingers move close to the palm whereas the thumb and index do not move. Finally, dynamic hand *gesture 4* involves motion of the index, middle, ring and little fingers moving close to the palm whereas the thumb remains stationary. Fig. 6 illustrates two of these dynamic hand gestures.

## VI. RESULTS AND DISCUSSION

### A. Evaluation of MDTW

Here, we compare a particular hand gesture with itself, meaning the same vector describing how the fourteen finger joints evolve over successive time frames is compared to itself. When the comparison is completed, the minimum-distance value of a warping path obtained is 0°. This is a perfect result since the similarity is 100%.

### B. Comparing Different Gestures to a Reference Gesture

Here we designed 4 different hand gesture comparisons where in comparison 1 i.e. cf. #1, hand gestures 2, 3 and 4 are compared to hand gesture 1. In cf. #2, hand gestures 1, 3 and 4 are compared to hand gesture 2. In cf. #3, hand gestures 1, 2 and 4 are compared to hand gesture 3 and finally in cf. #4, hand gestures 1, 2 and 3 are compared to hand gesture 4.

From Fig. 7, hand gesture 1 and hand gesture 4 are similar since a minimum-distance of lowest value of 35.9° is obtained.

TABLE I  
FEATURE VECTOR IN THE EXPERIMENTAL SET-UP

Fingers of the Hand	MCP	PIP	DIP/IP
Thumb	$\theta_1$		$\theta_2$
Index	$\theta_3$	$\theta_4$	$\theta_5$
Middle	$\theta_6$	$\theta_7$	$\theta_8$
Ring	$\theta_9$	$\theta_{10}$	$\theta_{11}$
Little	$\theta_{12}$	$\theta_{13}$	$\theta_{14}$

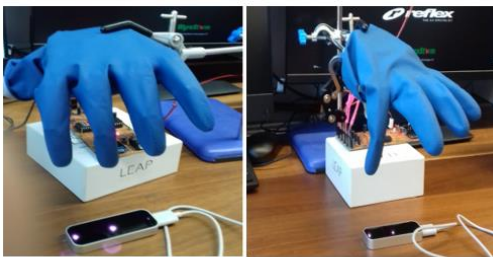
On the other hand, hand gesture 2 and hand gesture 3 experience the largest variation since a high value of minimum-distance of  $248.5^\circ$  is obtained. From our experimental design, this makes sense since hand gesture 1 is nearly identical to hand gesture 4 whereas hand gesture 2 can be regarded as the direct opposite to hand gesture 3.

### C. Comparing Gestures at Varied Distances from the LM

In this experimental setup we considered four different configurations i.e.  $D1$  is 5 cm to the right of the centre-line of the LM from the middle of the palm of the robotic hand whereas  $D2$  is 5 cm to the left.  $D3$  is 5 cm in front of the centre-line of the LM from the middle of the palm of the robotic hand whereas  $D4$  is 5 cm behind. During the execution of a dynamic hand gesture, we maintained a moderate speed, a vertical distance of 15 cm from the surface of the LM to the palm of the robotic hand, and the robotic hand was fixed in a horizontal position relative to the LM with its palm facing downwards.

For comparison purposes,  $D1D2$  implies the dynamic hand gesture performance at  $D1$  is compared with the dynamic hand gesture performance at  $D2$ . This notation is applied to the remaining cases as well, for example  $D3D4$  implies dynamic hand gesture performance at  $D3$  is compared with dynamic hand gesture performance at  $D4$ .

From Fig. 8, the minimum-distances range is from around  $5^\circ$  to  $12.5^\circ$ . These are quite low values compared to minimum-distance values obtained in Fig. 7 where we compared dynamic hand gestures to a specific gesture. Hence reasonable distance from the centre-line of the LM to the middle of the palm of a hand does not affect gesture performance.



Gesture 1

Gesture 2

Fig. 6. Some of the hand gestures used to facilitate implementation of MDTW.

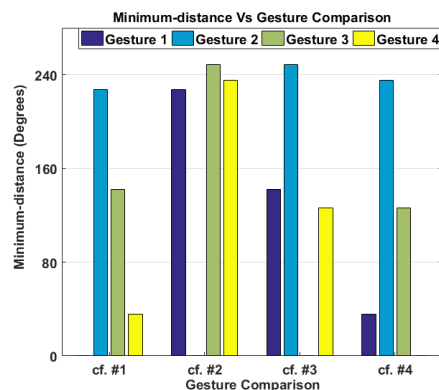


Fig. 7. Minimum-distance against gesture comparison.

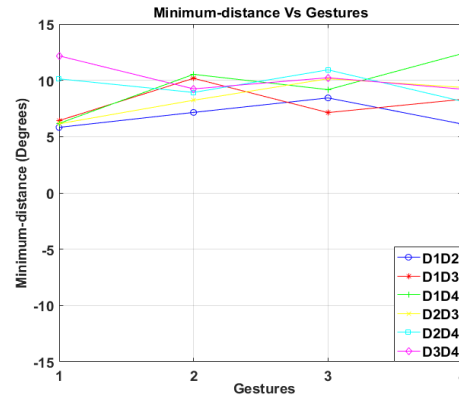


Fig. 8. Varied distances for hand gesture comparisons.

### D. Comparing Gestures at Varied Speeds

In this experimental setup, we arranged four different configurations, i.e.  $S1$  is the baseline case, where a dynamic hand gesture is performed at a slow speed,  $S2$  is when a dynamic hand gesture is performed at a speed twice as fast as a baseline one,  $S3$  is when a dynamic hand gesture is performed at a speed three times faster than the baseline case and  $S4$  is when a dynamic hand gesture is performed at a speed four times faster than the baseline one.

For comparison purposes,  $S1S2$  implies the dynamic hand gesture performance at  $S1$  is compared with the dynamic hand gesture performance at  $S2$ . This notation applies to the remainder of the experimental setup, for example,  $S3S4$  implies dynamic hand gesture performance at  $S3$  is compared with dynamic hand gesture performance at  $S4$ . During a dynamic hand gesture evaluation we placed the robotic hand at the LM centre-line, maintaining a vertical distance of 15 cm from the surface of the LM to the palm of the robotic hand, and the robotic hand was fixed in a horizontal position with its palm facing downwards.

As illustrated in Fig. 9, the minimum-distances range from approximately  $3^\circ$  to  $11^\circ$ . It is important to note that these values are in the same range as those obtained when comparisons are made based on distances from the LM to the robotic hand (Fig. 8). This implies variable speed of hand gestures does not significantly affect gesture comparison performance.

### E. Comparing Gestures when the Robotic Hand is Rotated

In this setup we rotate the robotic hand while performing dynamic hand gestures considering four different scenarios, i.e.  $R1$ ,  $R2$ ,  $R3$ , and  $R4$ .  $R1$  is when the palm of the robotic hand is in horizontal position and facing downwards.  $R2$  is when the robotic hand is roughly rotated at  $30^\circ$  and its thumb facing upwards.  $R3$  is when the robotic hand is roughly rotated at  $60^\circ$  and its thumb facing upwards.  $R4$  is when the robotic hand is in a vertical position and its thumb facing upwards. All these rotations are relative to the surface of LM. During this dynamic hand gesture evaluation, we placed the robotic hand at LM centre-line, maintained a vertical distance of 15 cm from the surface of the LM to the palm of the robotic hand, and maintained a moderate speed.

For comparison purposes, *R1R2* implies the dynamic hand gesture performance at *R1* is compared with the dynamic hand gesture performance at *R2*. This notation applies to the remaining cases as well, for example *R3R4* implies dynamic hand gesture performance at *R3* is compared with dynamic hand gesture performance at *R4*.

As shown in Fig. 10, the minimum-distances range from approximately  $4^\circ$  to  $14^\circ$ . These minimum-distance values are in the same range as those in Fig. 8 and Fig. 9. However, it is important to note that the *R1R2* comparison registers the lowest minimum-distance values. On the other hand, *R1R4* registers the highest minimum-distance values. This may suggest that a physiotherapist should encourage patients to avoid unnecessary rotations of their hands.

From the preceding analysis, regarding comparisons between gestures that are supposedly performed by patients with those that would be recommended by a physiotherapist, we observe that there is no concern about requiring a patient to maintain a particular horizontal distance from the LM to their hand for sensible rotations of the hand, and varying speed of gestures performed by different patients with various kinds of hand related injuries, as these factors do not significantly affect the gesture comparison. However, it is important to note that we maintained vertical distances around 15 cm from the surface of the LM to the palm of the robotic hand in all our experiments due to our previous work [17], where we established that accurate LM readings are obtained around this height.

#### F. Detailed Comparison for Hand Gestures

Here we undertake a comparison of all the four dynamic hand gestures performed in all the scenarios we explained in subsections C through to E i.e. at various distances, at various speeds, and at various rotations. For each of the four dynamic hand gestures, both query and reference gestures were at *D1*, *D2*, *D3*, *D4*, *S1*, *S2*, *S3*, *S4*, *R1*, *R2*, *R3*, and *R4*.

For gesture 1, the lowest minimum-distance value is  $0^\circ$  and the highest value is  $12.8^\circ$ . For gesture 2, it was observed that the minimum-distance values range from  $0^\circ$  to  $18.2^\circ$ . For gesture 3, the minimum-distance values range from  $0^\circ$  to  $17.4^\circ$  and for gesture 4, the minimum-distance values range from  $0^\circ$  to  $18.8^\circ$ .  $0^\circ$  was obtained when a dynamic hand gesture was compared to itself in the same scenario under the same conditions. This is expected since the similarity is 100%. The same scenario same condition situations were when, for example, a query at *D1* was matched with a reference at *D1*, or a query at *R1* was compared with a reference at *R1*, etc. Furthermore, minimum-distance values are lower compared to those obtained when comparison is made between different hand gestures. This signifies the acceptable distance from the LM to the hand and indifference to speed of hand gestures during rehabilitation hand exercises. Sensible rotations of the hand also have little impact on hand gesture comparison performance.

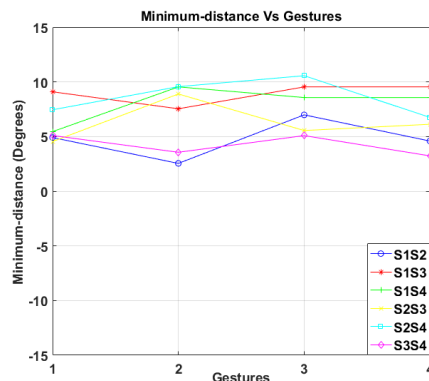


Fig. 9. Varied speeds for hand gesture comparisons.

#### G. Experimental Limitations

Although the four different hand rehabilitation dynamic gestures were sufficient for the implementation and evaluation of MDTW, it could have been a good idea to extend the number of gestures. In this study we were constrained by the robotic hand performing only flexing and extending finger gestures. It was unable to execute complex hand gestures like abduction and adduction of fingers. Further, there was no need to recruit patients with hand related injuries like those recovering from stroke though if patient evaluation was considered could have added more significant benefits in the field of hand therapy.

#### VII. CONCLUSION AND FUTURE WORK

We have proposed, implemented and evaluated Multi-dimensional Dynamic Time Warping (MDTW) to establish how similar or different a query dynamic hand gesture could be in relation to a reference dynamic hand gesture. The approach is robust in the way that it can distinguish quite different hand dynamic gestures by yielding a high value of a minimum-distance metric. Furthermore, the method can produce a low value of a minimum-distance if the hand dynamic gestures under comparison are closely similar. Therefore, in the context of dynamic rehabilitation hand gestures performed by a patient and compared with reference gestures recommended by a physiotherapist, meaningful feedback can be generated concerning how well the patient is mimicking the reference gestures.

In the future we shall extend on the kinds of hand gestures by possibly improvise another different robotic hand that can perform abduction and adduction of fingers gestures. It would also be relevant to recruit both patients with hand related injuries and healthy individuals and evaluate how the two groups perform hand rehabilitation gestures. Obtained results can help a physiotherapist generate a more realistic feedback to the patient. Finally, since the recognition rate in LM could slightly vary depending on the speed of hand gestures, this can also be an interesting area for future research.

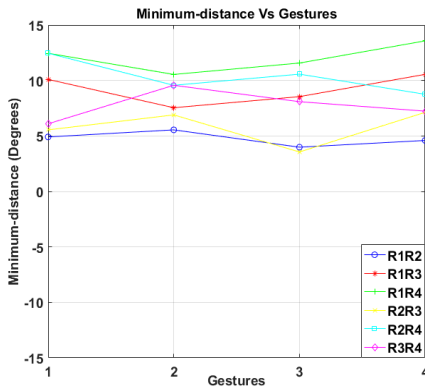


Fig. 10. Varied rotations for hand gesture comparisons.

## REFERENCES

- [1] D. Bercht, T. Boisvert, J. Lowe, K. Steams, and A. Ganz, "ARhT: A portable hand therapy system," in Proc. Int. Conf. IEEE Eng. Med. Biol. Soc., pp. 264-267, Nov. 2012.
- [2] K. M. Vamsikrishna, D. P. Dogra, and M. S. Desarkar, "Computer-vision-assisted palm rehabilitation with supervised learning," IEEE Trans. Biomed. Eng., vol. 63, no. 5, pp. 991-1001, May 2016.
- [3] D. Wade and B. Jong, "Recent advances in rehabilitation," Biomed. J., vol. 320, pp. 1385-1388, May 2000.
- [4] H. Walugembe, C. Phillips, J. Requena-Carrion, and T. Timotijevic, "Gesture recognition in Leap Motion using LDA and SVM," in Proc. IEEE Int. Conf. Comp. Elec. Comms. Eng., pp. 56-61, Aug. 2019.
- [5] C. B. Redd, and S. J. Morris, "A wireless sensory feedback device for real-time gait feedback and training," IEEE ASME Trans Mechatron, vol. 17, no. 3, pp. 425-433, June 2012.
- [6] L. E. Sucar, R. Luis, R. Leder, J. Hernandez and I. Sanchez, "Gesture therapy: A vision-based system for upper extremity stroke rehabilitation," in Proc. Int. Conf. of the IEEE EMBS, pp. 3690-3693, Sept. 2010.
- [7] J. Zariffa, and J. D. Steeves, "Computer vision-based classification of hand grip variations in neurorehabilitation," in Proc. IEEE Int. Conf. on Rehab Robot., pp. 1-4, July 2011.
- [8] M. Naeemabadi, B. Dinesen, O. K. Andersen, and J. Hansen, "Influence of a marker-based motion capture system on the performance of Microsoft Kinect v2 skeleton algorithm," IEEE Sens. J., vol. 19, no. 1, pp. 171-179, Jan. 2019.
- [9] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using Kinect sensor," IEEE T Multimedia, vol. 15, no. 5, pp. 1110-1120, Aug. 2013.
- [10] C. Wang, Z. Liu, and S. Chan, "Superpixel-based hand gesture recognition with Kinect depth camera," IEEE T Multimedia, vol. 17, no. 1, pp. 29-39, Jan. 2015.
- [11] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of Leap Motion controller," Sensors, vol. 13, no. 5, pp. 6380-6393, May 2013.
- [12] T. S. Sarkar, S. Das, B. Chakraborty, and H. S. Dutta, "Absolute encoder-based dual axis Tilt sensor," IEEE Sens. J., vol. 19, no. 7, pp. 2474-2481, April 2019.
- [13] M. Ghanbari and M. J. Yazdanpanah, "Delay compensation of Tilt sensors based on MEMS Accelerometer using data fusion technique," IEEE Sens. J., vol. 15, no. 3, pp. 1959-1966, Mar. 2015.
- [14] L. Ding et al., "Camera-based mirror visual feedback: potential to improve motor preparation in stroke patients," IEEE Trans. Neural Syst. and Rehabilitation Eng., vol. 26, no. 9, pp. 1897-1905, Sept. 2018.
- [15] M. Ye, C. Yang, V. Stankovic, L. Stankovic, and A. Kerr, "A depth camera motion analysis framework for telerehabilitation: motion capture and person-centric kinematics analysis," IEEE J-STSP, vol. 10, no. 5, pp. 877-887, Aug. 2016.
- [16] B. J. Ranger et al., "3D Ultrasound imaging of residual limbs with camera-based motion compensation," IEEE Trans. Neural Syst. and Rehabilitation Eng., vol. 27, no. 2, pp. 207-217, Feb. 2019.
- [17] H. Walugembe, C. Phillips, J. Requena-Carrion, and T. Timotijevic, "Characterizing and compensating for errors in a Leap Motion using PCA," in Proc. IEEE Int. Signal Process. Info. Security (ICSPIS), pp. 1-4, Nov. 2018.
- [18] G. Yang et al., "An IoT-enabled stroke rehabilitation system based on smart wearable armband and machine learning," IEEE J. Transl. Eng. He., vol. 6, pp. 1-10, May 2018.
- [19] S. Salim, W. N. Zakaria, A. Nizhan, and M. M. A. Jamil, "Integration of tilt sensors as a device for monitoring rehabilitation process," in Proc. IEEE Int. Conf. Ctrl. Sys. Comp. Eng., pp. 232-235, Nov. 2014.
- [20] S. Makris et al., "Dual arm robot in cooperation with humans for flexible assembly," CIRP Annals – Mfg. Technol. 66, pp. 13-16, 2017.
- [21] A. Vaitkevicius et al., "Recognition of American sign language gestures in a virtual reality using leap motion," Appl. Sci., vol. 9, no. 3, pp. 1-16, Jan. 2019.
- [22] A. Bhardwaj, A. Grover, P. Saini, and M. Singh, "Contact dynamics emulation using leap motion controller," in Proc. Int. Conf. on Adv. in Comput. Data Sci., pp. 262-271, July 2017.
- [23] P. Tormene, T. Giorgino, S. Quaglioni, and M. Stefanelli, "Matching incomplete time-series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation," Artif. Intell. Med., vol. 45, no. 1, pp. 11-34, Jan. 2009.
- [24] A. L. Uribe-Hurtado, M. Orozco-Alzate, and E. A. Rodriguez-Soto, "Dynamic time warping dissimilarity matrices," IEEE Potentials, vol. 37, no. 5, pp. 34-42, Oct. 2018.
- [25] K. Mendhurwar, Q. Gu, S. Mudur, and T. Popa, "The discriminative power of shape an empirical study in time series matching," IEEE Trans. Vis. Comput. Graph., vol. 24, no. 5, pp. 1799-1813, May 2018.
- [26] M. Lu, and H. Lin, "Parallel algorithms for the longest common subsequence problem," IEEE Trans. Parallel Distrib. Syst., vol. 5, no. 8, pp. 835-848, Aug. 1994.
- [27] J. Herranz, J. Nin, and M. Sole, "Optimal symbol alignment distance: a new distance for sequence of symbols," IEEE Trans. Knowl. Data Eng., vol. 23, no. 10, pp. 1541-1554, Oct. 2011.
- [28] P. Marteau, "Time warp edit distance with stiffness adjustment for time series matching," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 2, pp. 306-318, Feb. 2009.
- [29] D. Jin, R. Li, and J. Xu, "Multiscale community detection in functional brain networks constructed using dynamic time warping," IEEE Trans. Neural Syst. Rehabilitation Eng., vol. 28, no. 1, pp. 52-61, Jan. 2020.
- [30] M. E. Munich, and P. Perona, "Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification," in Proc. IEEE Int. Conf. Comp. Vis., pp. 1-8, Aug. 2002.
- [31] D. Liu et al., "Downsizing without downgrading: approximated dynamic time warping on nonvolatile memories," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 39, no. 1, pp. 131-144, Jan. 2020.
- [32] A. C. S. Douglass and J. B. Harley, "Dynamic time warping temperature compensation for guided wave structural health monitoring," IEEE Trans. Ultrason. Ferroelectr., Freq. Control, vol. 65, no. 5, pp. 851-861, May 2018.
- [33] R. Yang et al., "Mechanical fault diagnostics of power transformer on-load tap changers using dynamic time warping," IEEE Trans. Instrum. Meas., vol. 68, no. 9, pp. 3119-3127, Sept. 2019.
- [34] G. Ponraj and H. Ren, "Sensor fusion of Leap Motion controller and Flex Sensors using Kalman filter for human finger tracking," IEEE Sens. J., vol. 18, no. 5, pp. 2042-2049, Mar. 2018.
- [35] American society for surgery of the hand. (2019, Nov. 19). [Online]. Available: <https://www.assh.org/handcare/Anatomy/Joints>
- [36] M. Ibrahim and D. Mulvaney, "Geometrical-based lip-reading using template probabilistic multi-dimensional dynamic time warping," J. Vis. Commun. Image Represent., vol. 30, pp. 219-233, Jul. 2015.
- [37] S. Elsworth, "Tutorial on Dynamic Time Warping, School of Mathematics," University of Manchester, Sept. 2017.
- [38] W. Lu, Z. Tong, and J. Chu, "Dynamic hand gesture recognition with leap motion controller," IEEE Signal Process. Lett., vol. 23, no. 9, pp. 1188-1192, Sept. 2016.
- [39] S. Rautaray, and A. Agrawal, "Interaction with virtual game through hand gesture recognition," in Proc. IEEE Int. Multimedia, Signal Process. Comm. Technol., pp.244-247, Dec. 2011.
- [40] E. Ohn-Bar, and M. M. Trivedi, "Hand gesture recognition in real time for automatic interfaces: a multimodal vision-based approach and evaluations," IEEE Trans. Intell. Transp. Syst., vol. 15, no. 6, pp. 2368-2377, Dec. 2014.
- [41] J. Vinstrup et al., "Hand strengthening exercises in chronic stroke patients: dose-response evaluation using electromyography," J. Hand Ther., pp. 111-121, May 2017.