Medical gesture recognition using dynamic arc length warping

Jenny Cifuentes, Minh Tu Pham, Richard Moreau, Pierre Boulanger, Flavio Prieto

Medical Gesture Recognition Using Dynamic Arc Length Warping

Jenny Cifuentes [1], Minh Tu Pham [2], Richard Moreau [3], Pierre Boulanger [4], Flavio Prieto [5]

Program of Electrical Engineering - Universidad de la Salle [1]

Laboratory Ampere - INSA de Lyon [2,3]

Computing Sciences Department- University of Alberta [4]

Mechanic and Mechatronic Department - Universidad Nacional de Colombia [1,5]

Abstract

Hand gesture recognition is a promising research area often used in applications of human-computer interactions in the medical field. In this paper, we present a novel approach to differentiate gestures based on an arc-length parametrization and a curvature analysis of 3D trajectories. This new method called Dynamic Arc Length Warping (DALW) can outperform classic Multi Dimensional-Dynamic Time Warping (MD-DTW) algorithm as it is truly invariant to sensor location and more tolerant to temporal distortions. Experimental validation of the algorithm is presented using different gestures and sensors in biomedical applications: an exoskeleton apparatus,

[1]jacifuentesq@gmail.com, Tel: +57 3166243431
[2]minh-tu.pham@insa-lyon.fr
[3]richard.moreau@insa-lyon.fr
[4]pierreb@ualberta.ca
[5]faprietoo@unal.edu.co

surgical gestures captured by an instrumented laparoscopic device and finally, a birth simulator with an instrumented forceps. A basic perceptron multilayer neural network was implemented in order to perform the classification. Results involve an average increase of 7.14% in the classification rates by using DALW distance, compared to the classical MD-DTW.
Keywords: Gesture Classification; Curvature Analysis; Dynamic Arc Length Warping; Hand Motion Tracking.

---

1. Introduction

Accurate hand gesture detection and tracking have been a challenging problem with medical applications in human-computer interfaces, motion capture, and scene understanding [35, 19]. In fact, the latest advances, in the fields of pattern recognition and computer vision, have made that hand gesture classification approach becomes an option to assess competences in simulation-based environments [18, 13].

Different strategies have been proposed to acquire the relevant information of gestures recognition systems. Some approaches use hardware devices including vision systems, torque and force sensors, or EMG electrodes in order to facilitate the extraction of comprehensible features, such as kinematic, dynamic or electrical variables to characterize each gesture [27, 38, 36].

In order to carry out the comparison, a local analysis has been proposed on numerous occasions to provide a qualitative description [24, 8]. In this particular analysis, the complete medical procedure is decomposed and seg-

mented into smaller units. Under such approach, measures of motor performance are recorded including kinematic and kinetic measurements, such as position, velocity, acceleration, and force/torque values [25, 9]. Results, obtained in this field, have shown, in both open and MIS (Minimally Invasive Surgery) simulators, a correlation between movements made and objective skill measurements in simulation (Spearman coefficient 0.53) [6, 7]. Based on this idea, force/torque recordings have differentiated trajectories performed by novice from expert surgeons in a porcine model of MIS [29]. In this line of thought, the methodology developed in [30], based on a subset of hidden Markov modeling, has been proposed. The force/torque measurements were used to specify force/torque signatures associated to different categories of tool/tissue interactions, resulting in a surgical performance index which represents a ratio of statistical similarity among different surgeons. It is important to note that results associated to this method, require large data sets to be decomposed manually. In [31], on the other hand, the acquisition, storage, analysis, and classification are based on kinematic and dynamic data of surgeon's hand postures during a simulated operation.

However, some problems have been found using sensor data (position, speed, strength, etc) during the classification process. One of them is based on data synchronization, because signals involved could have different sizes. In this case, a point by point comparison may provide unreliable results. In order to overcome this problem, several strategies have been proposed in order to shrink or expand data along the time axis. In particular, ap-

proaches like Dynamic Time Warping (DTW) [17, 1] or Longest Common Subsequence (LCSS) [14, 34] can compute the optimal alignment between 2 sequences. Numerous researches, using these strategies, have been developed in the area of gesture recognition, describing their advantages compared to classical recognition approaches [28, 3]. As a case in point, Pham *et al* [26] use DTW to match and compare the curvature of two 3D trajectories.

However, these strategies to synchronize and analyze trajectories from surgical movements include a time parametrization for each path [12]. Consequently, this idea implies that gestures are not time invariant, and paths that follow identical trajectories at different velocities have different characteristics. Additionally, 3D position coordinates are not invariant under rotation and scaling transformations which could cause errors during the synchronization.

The approach proposed in this paper includes a data parametrization of position trajectories, taking into account the cumulative arc-length. In particular, transformation to cumulative arc-length parametrization does not exclude information in comparison with the time-domain representation of a signal, since this conversion is analogous to a time normalization operation. Specifically, this parametrization describes a traversal at unit-speed of every trajectory [10]. This fact supports the main advantages of this strategy, based on the alternative to compare similar trajectories, carried out at different velocities, and analyzing the corresponding movements independently of a sensor coordinate system. Finally, a curvature calculation is performed in

order to carry out the gesture analysis and recognition, which is invariant under sensor coordinate systems. In addition, this method allows to reduce, as well, the dimensionality of the problem from 3D to 1D. This reduction, in particular, can simplify the classification problem and makes easier to implement and understand. The experimental results are focused on the acquisition and classification of diverse medical gestures using three different devices.

This paper is organized as follows. Section 2 presents the mathematical foundation of the algorithm. Section 3 describes the algorithm in detail. Section 4 reports experimental results and a comparison to classical dynamic time warping technique. And finally, in Section 5, we conclude and propose future directions.

2. Mathematical Background

In this section, the mathematical basis of the proposed approach is presented. The emphasis is on 3D-trajectories transformation to a space-time invariant signal, used to obtain the comparison between different gestures.

2.1. Time Series with Cumulative Arc-length Parametrization

A 4D-trajectory P is defined by (x, y, z, t), represented as a vector by $\vec{r}(t) = [x(t),\ y(t), z(t)]$ in 3D-space where time $t \in [a, b]$ is then a dependent parametric variable. The difficulty of classical time parametrization is that movements are not time-dependent as for example a circular gesture is

basically independent of the rotation speed. A different way to parametrize space, which is independent of coordinate system and time, is using the concept of cumulative arc-length. This variable depends on the total length L of each 3D-trajectory from the beginning to the end of the gesture $[t_a, t_b]$ and is specified by:

$$S = \int_{t_a}^{t_b} ||\vec{r}'(t)||dt, \tag{1}$$

where $\vec{r}'(t)$ is the first derivative of the signal with respect to time. P can be reparameterized by using a normalized parameter s called the cumulative arc-length s, designated by:

$$s = \frac{1}{S} \int_{t_a}^{t} ||\vec{r}'(t)||dt, \tag{2}$$

where $s \in [0, 1]$. The parametrization of each signal, using time, means that movements are not time invariant, and trajectories that follow identical paths at different speeds are substantially different. Parametrization, proposed in this work, overcomes this problem, using the cumulative arc length information, which characterizes the curve shape and does not depend on a particular coordinate system.

2.2. Trajectory Curvature

Given a point, the curvature vector $\vec{r}''(s)$ of a trajectory is determined as the rate of change of the unit tangent vector with respect to cumulated

arc-length. Curvature κ, on that premise, is defined as the length of the curvature vector as follows:

$$\kappa(s) = \|\vec{r}''(s)\|. \tag{3}$$

To be specific, the curvature, at a given point, in a trajectory is a measure of how fast the curve modifies the direction at that point. However, when parametrization using the cumulative arc-length is implemented, final trajectory is sampled at unequal intervals in space (see Figure 1). Consequently, classical techniques of numerical differentiation cannot be applied to calculate the curvature.

One way to overcome this problem is fitting the data using polynomial functions. Under this approach, a second-order Lagrange polynomial interpolation strategy is implemented to fit each group of 3 adjacent points. This interpolation technique is selected in this work because it has the following advantages:

1. Simplicity: The Lagrange interpolating polynomial is computed without solving a simultaneous-equations system;
2. Roundoff: The Lagrange interpolating polynomial is more tolerant to roundoff compared to other polynomial interpolation techniques;
3. Unequally Spaced Data: The most important reason of using this strategy is that it does not require evenly spaced sampled data signals.

The second order polynomial can be analytically differentiated twice using

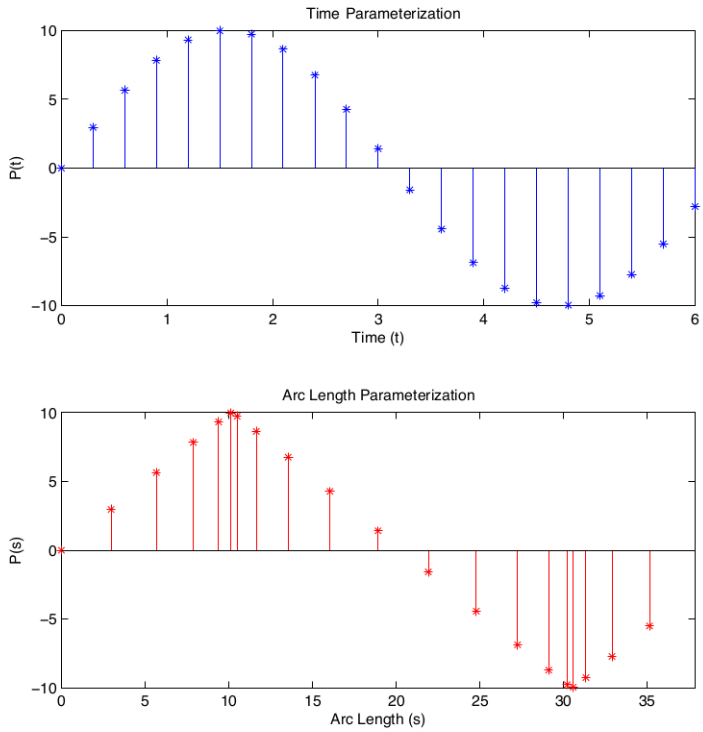

Figure 1: Arc-length and time parameterization

the following procedure: firstly, a Lagrange interpolating polynomial is used to fit each group of 3 adjacent points:

$$f_n(s) = L_{i-1}(s)p_n(s_{i-1}) + L_i(s)p_n(s_i) + L_{i+1}(s)p_n(s_{i+1}), \tag{4}$$

$$\begin{aligned} f_n(s) =& \frac{(s-s_i)(s-s_{i+1})}{(s_{i-1}-s_i)(s_{i-1}-s_{i+1})}p_n(s_{i-1}) + \frac{(s-s_{i-1})(s-s_{i+1})}{(s_i-s_{i-1})(s_i-s_{i+1})}p_n(s_i) \\ &+ \frac{(s-s_{i-1})(s-s_i)}{(s_{i+1}-s_{i-1})(s_{i+1}-s_i)}p_n(s_{i+1}), \end{aligned} \tag{5}$$

where f_n is denoted as the Lagrange polynomial function for each coordinate x, y, and z, respectively. L_i, on the other hand, is the Lagrange basis function, and p_n are the points to interpolate for each axis. The Lagrange polynomial is differentiated twice for the purpose of calculating the curvature values of each trajectory:

$$f_n'(s) = \frac{(s-s_i)+(s-s_{i+1})}{(s_{i-1}-s_i)(s_{i-1}-s_{i+1})}p_n(s_{i-1}) + \frac{(s-s_{i-1})+(s-s_{i+1})}{(s_i-s_{i-1})(s_i-s_{i+1})}p_n(s_i) + \frac{(s-s_{i-1})+(s-s_i)}{(s_{i+1}-s_{i-1})(s_{i+1}-s_i)}p_n(s_{i+1}), \tag{6}$$

$$f_n'(s) = \frac{2s-s_i-s_{i+1}}{(s_{i-1}-s_i)(s_{i-1}-s_{i+1})}p_n(s_{i-1}) + \frac{2s-s_{i-1}-s_{i+1}}{(s_i-s_{i-1})(s_i-s_{i+1})}p_n(s_i) + \frac{2s-s_{i-1}-s_i}{(s_{i+1}-s_{i-1})(s_{i+1}-s_i)}p_n(s_{i+1}), \tag{7}$$

$$f_n''(s) = \frac{2p_n(s_{i-1})}{(s_{i-1}-s_i)(s_{i-1}-s_{i+1})} + \frac{2p_n(s_i)}{(s_i-s_{i-1})(s_i-s_{i+1})} + \frac{2p_n(s_{i+1})}{(s_{i+1}-s_{i-1})(s_{i+1}-s_i)}, \tag{8}$$

where s_{i-1}, s_i, s_{i+1} is the cumulative arc length for 3 consecutive points.

2.3. Filtering Process

A low-pass filter, without phase shift and without magnitude distortion, is required in order to reduce the noise. This filter is implemented using a

non-causal zero-phase digital filter, which processes the input data using an IIR low-pass Butterworth filter in both, forward and reverse direction.

The cut-off frequency ω_c of the low-pass filter must be set to prevent the magnitude distortion on the filtered signals within the trajectory bandwidth. A second filter is implemented to remove the higher frequency noise. Its cut-off frequency is determined experimentally by analyzing the FFT (Fast Fourier Transform) of each component of trajectories. Through this analysis, the frequency for which the magnitude values are less than 15% of the maximum, is computed (see Figure 2). The maximum value of these frequencies is used in order to avoid missing significative information for each coordinate. It is important to note that, in practice, for a real-time application, an online filtering technique should be taking into account [15, 16].

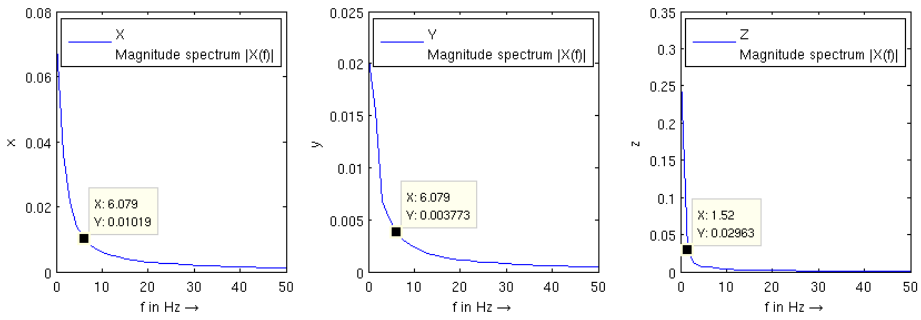

Figure 2: Fourier transform of each coordinate (Example).

3. Dynamic Arc-Length Warping Algorithm

Based on the formulation described in the previous section, numerical derivatives of two trajectories $\vec{r}_1(s)$ and $\vec{r}_2(s')$ are computed in order to calcu-

late the corresponding curvatures. From this perspective, the aim of DALW is to compare two curvature sequences $\kappa_1 = \{\kappa_1(s_1), \kappa_1(s_2), \kappa_1(s_3), ..., \kappa_1(s_l)\}$ of length $l \in \mathbb{N}$ and $\kappa_2 = \{\kappa_2(s'_1), \kappa_2(s'_2), \kappa_2(s'_3), ..., \kappa_2(s'_m)\}$ of length $m \in \mathbb{N}$. Based on this idea, a local dissimilarity matrix $\mathbf{F}$ is defined. The matrix elements measure the distance between the curvature values $\kappa_1(s_i)$ and $\kappa_2(s'_j)$, according to a chosen norm. In this case, the Euclidean norm is used for convenience. This local similarity metric is defined between any pair of elements $\kappa_1(s_i)$ and $\kappa_2(s'_j)$ using the following expression:

$$F(i,j) = (\kappa_1(s_i) - \kappa_2(s'_j))^2 \geq 0 \tag{9}$$

In general, $F(i,j)$ is small (low cost) if $\kappa_1(s_i)$ and $\kappa_2(s'_j)$ are similar to each other and otherwise $F(i,j)$ is large (high cost). Calculating the local cost measure between two sequences κ_1 and κ_2 conducts to a local dissimilarity matrix by $\mathbf{F}$. This process provides a basis to find an alignment between κ_1 and κ_2 by minimizing the overall cost function. A warping path of a curvature κ_n is defined by $\phi_n(k),\ k = 1, 2, ..., W$ (Equation 10).

$$\kappa'_n(s_j) = \kappa_n(s_{\phi(k)}), \tag{10}$$

where κ'_n is a new curvature array, W is the length of the warping path, and the s_j value corresponds to the previous signature value at s_k.

Using this notation, the warping functions $\phi_1(k)$ and $\phi_2(k)$ re-map the cumulative arc-length index of κ_1 and κ_2, respectively. Given ϕ, the total distance between the warped arc-length series can be calculated as:

$$d(k) = \sum_{k=1}^{W} \frac{1}{\alpha_{\phi(k)}} \Big(\kappa_1(s_{\phi_1(k)}) - \kappa_2(s_{\phi_2(k)}) \Big)^2 M_\phi \tag{11}$$

where $\alpha_{\phi(k)}$ is a per-step weighting coefficient and M_ϕ is a normalization constant, which allows distances to be comparable along different paths.

Regarding to per-step weighting coefficient α_ϕ, it allows to include different non-negative weights to vertical, horizontal, and diagonal directions based on particular preferences. In this proposed work, the weighting functions introduced in [32], were used (Equation 12). This function allows an equal preference for alignments, both in vertical and horizontal direction, which is also higher than the preference for an alignment in the diagonal direction.

$$\alpha_\phi(k) = \phi_1(k) - \phi_1(k-1) + \phi_2(k) - \phi_2(k-1) \tag{12}$$

The value of M_ϕ, on the other side, depends on the application and in most cases, it is defined as the length of the path, but it can also be omitted [33].

The optimal warping corresponds to the warping $\phi_1(k^*)$ and $\phi_2(k^*)$ that

minimizes the distance:

$$DALW(\kappa_1, \kappa_2) = d(k^*), \tag{13}$$

where k^* describes the index of the computed optimal warping path.

This result is equivalent to the similarity measure calculated between paths using DALW. This optimization problem is successfully solved by means of a dynamic programming technique [2].

Usually, constraints are imposed on the warping function $\phi_n(k)$ in order to ensure consistent optimal paths:

1. Boundary Condition: This criterion controls the beginning and the ending points of each trajectory.
   Beginning point : $\phi_1(1) = \phi_2(1) = 1$
   Ending point : $\phi_1(W) = l$; $\phi_2(W) = m$
2. Monotonicity Condition: The order of measurements acquired for each variable has a significative importance to the meaning of arc length. For this reason, it is required to impose a suitable monotonicity-constraint to keep the respective order in the arc length:
   $\phi_1(k+1) \geq \phi_1(k)$ and $\phi_2(k+1) \geq \phi_2(k)$
3. Step size Condition: This requirement restricts the warping path from long jumps (shifts in arc length axis) while signals are aligned. Basic step size condition is used and is formulated as:
   $|\phi_1(k+1) - \phi_1(k)| \leq 1$ and $|\phi_2(k+1) - \phi_2(k)| \leq 1$

The proposed approach (summarized in Figure 3) computes the arc length dynamic warping to match 2 curvature sequences. Using the information, associated to the alignment between each pair of points, the reconstruction between the original 3D-trajectories is carried out.

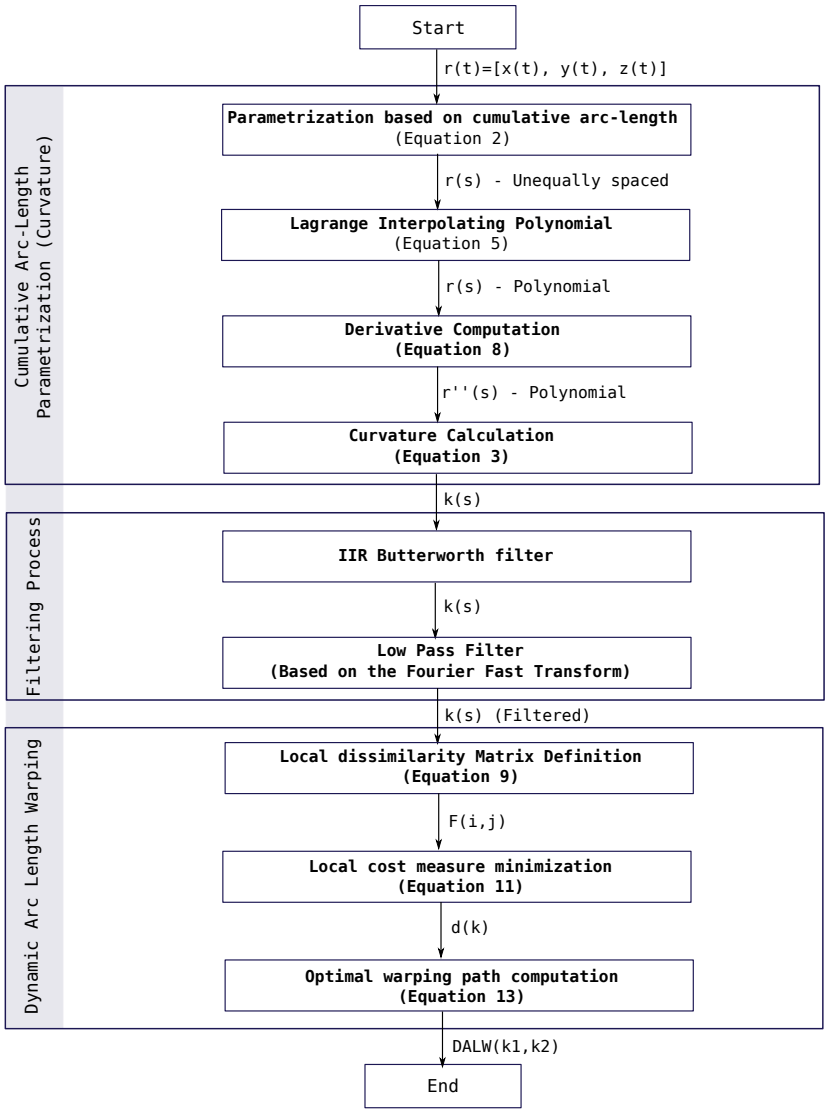

Figure 3: Dynamic Arc Lenth Warping Strategy.

DALW time complexity is $O(mn)$, which is the same as classical DTW, compared to $O(3mn)$ for MD-DTW. Empirically, both algorithms take approximately the same time. Previous works regarding to this topic have proposed a variety of optimizations for DTW [23]. Although, this discussion is not presented here for brevity, it is important to note that can be applied for DALW as well.

4. Experimental Results

The strategy, proposed in this paper, has been implemented using different 3D artificial trajectories, obtaining not only efficient quantitative but also qualitative results in comparison with other techniques such as Multi-Dimensional Dynamic Time Warping MD-DTW [4]. The objective of this section is to develop a validation of this method using real generic human hand gestures (Case 4.1 and 4.3) and real surgical gestures (Case 4.2) based on position data information.

4.1. Gesture Comparison Using an Exoskeleton

In order to test the algorithm proposed in this work, basic gestures were analyzed. The position for each trajectory was recorded by means of an upper extremity exoskeleton [22]. Figure 4 presents the exoskeleton employed and its kinematic model. The device has 4 degrees of freedom: 3 for the shoulder (internal-external rotation, abduction-adduction, flexion-extension) and 1 for the elbow (flexion-extension).

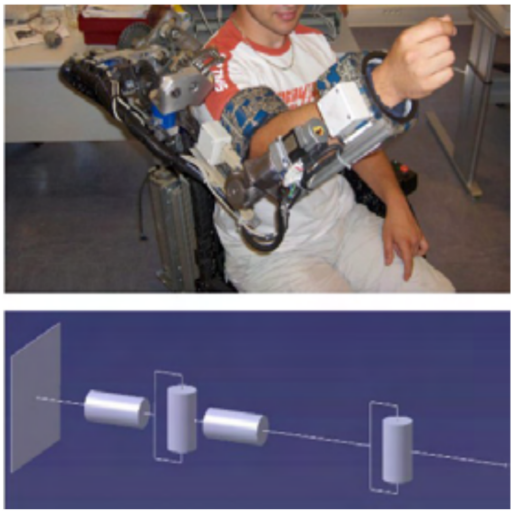

Figure 4: The upper extremity exoskeleton.

In this work, a total of 10 participants were involved. They were asked to do two different gestures (10 times each) with their right arm (all subjects are right handed). The first trajectory was a communication gesture as the participant wanted to say stop to someone running towards him/her. The second one was a basic gesture where the participant had to locate his/her hand as there was a wall in front of him/her. These gestures were chosen because of their similarity from a kinematic viewpoint. Figure 5 presents an example of position trajectories for each gesture. This figure illustrates the motion of a subject's wrist.

To avoid disturbances for each subject during the experiment, a gravity compensation technique was implemented to provide the illusion that the exoskeleton had no weight [21]. Using this device, the displacement of the subject's right arm was recorded and gestures were decomposed in different

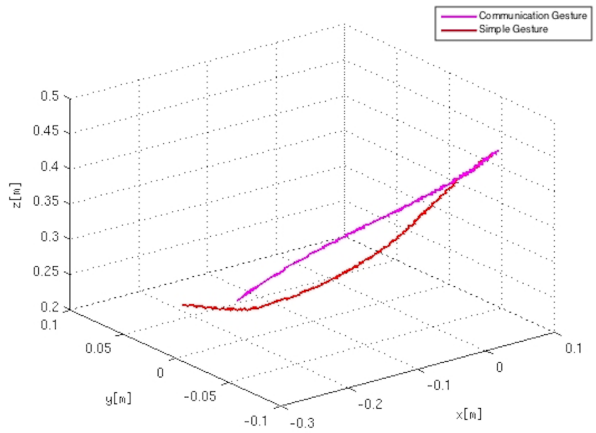

Figure 5: Position trajectory for different gestures.

anatomic planes for further studies on human joints. After the acquisition process of position trajectories, a filtering stage, described in section 2.3, was implemented.

In order to evaluate the significance of the data, previously described, a statistical test with a significance level of 5 % was performed. Because of the sample size, it is not possible to conclude if the data are normally distributed. For this reason, a nonparametric test was chosen to analyze the measurements. To be more specific, a Wilcoxon Signed-Rank test was used to compute the difference between each set of data and analyze these differences [37]. The Wilcoxon test is then used to test the null hypothesis that two data sets have the same continuous distribution. In this way, DALW and DTW distances were calculated between the same and different trajectories, and the results associated were compared in order to evaluate if data are significantly different. The p-values for DALW and DTW distances, calculated using the

comparison between trajectories performed by the same person and different participants, are 0.0145 and 0.0202, respectively. The results (p-value< 0.05) show that there is a significant difference between DALW and DTW distances computed for trajectories associated to only one participant and to different people. As a result, each pair of trajectories were aligned according to four different analysis (Table 1).

Table 1: Different Experiments.

Nature of Gestures	**Carried out by**
Same	Same person
	Different people
Different	Same person
	Different people

DTW and DALW distances were computed based on position data following these 4 criteria. Results associated to DTW and DALW similarity values for each experiment, based on position data, can be seen in Figure 6.

Results show the mean, maximum and minimum values of the similarity measurements, representing the average and range of acquired data, based on the analysis described in Table 1. Figure 6 presents distances obtained with both algorithms (DALW and DTW). It is possible to differentiate the same and different gestures regardless of whether these trajectories were performed by one participant or by different people. Specifically, deviations generated using the DALW algorithm are smaller than those obtained using the DTW algorithm, in most cases. These results show it is possible to differ-

entiate among the gestures proposed in this experimental setup. In addition, Figure 6 shows that, based on the fact that trajectories, associated to each gesture, are closely bounded to the people who execute them, it is possible to get smaller distances, for the experiment involving different gestures performed by the same participant, compared to the experiments involving similar gestures performed by different people.

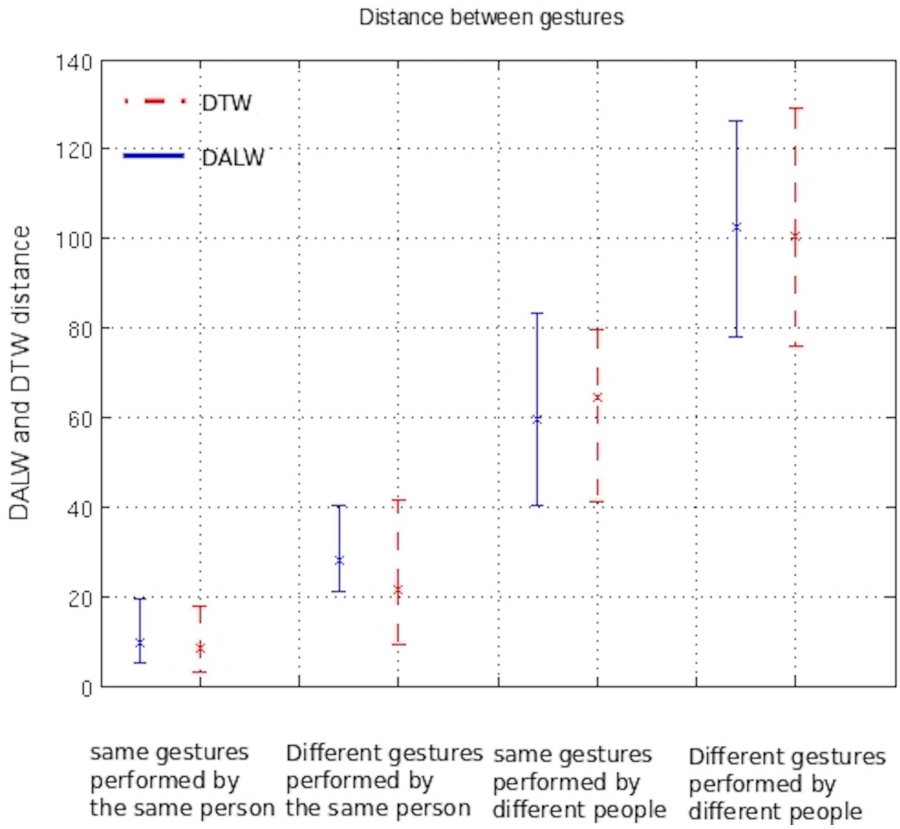

Figure 6: DTW and DALW distance computed using position data

In order to evaluate the advantages of DALW distances, a basic perceptron multilayer neural network was implemented. In this case, the structure

is defined by an input, associated to the DTW and DALW distances, 10 neurons in a hidden layer, and 4 neurons in the output layer, relating the different cases summarized in Table 1. Classification rates, associated to the different 4 experiments, were 88.1 % for DTW distances and 96.8 % for DALW distances, which implies a significative improvement by using the approach proposed in this work.

4.2. Gesture Comparison Using an Instrumented Laparoscopic Device

This section describes the surgical gestures evaluation using a laparoscopic training system. Figure 7 shows an instrument prototype that includes a Yaw-Roll actuated tip, which was employed to record different gestures. Handle and shaft orientations were decoupled by means of a free ball joint to improve the ergonomic performance. This disposition allows to explore the whole intra-abdominal workspace and prevent excessive wrist flexion or deviation. The robotic instrument configuration that was established is the Standard Fixed Handle Configuration, where joints were locked in their central position [11].

The training system, used in this part of the study, allows to implement a characteristic configuration of a real surgery where the instrument movements remain confined in a small region of the intra-abdominal workspace. Five participants without experience were involved in this experiments and five repetitions were performed for a particular pick-and-place task. Each experiment included one of five positions in the virtual abdomen to cover the

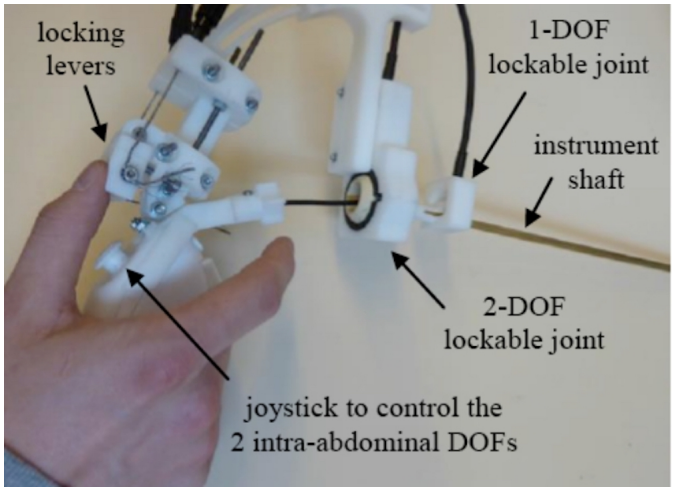

Figure 7: Laparoscopic Device.

whole workspace, taking 5 repetitions for each gesture (Figure 8).

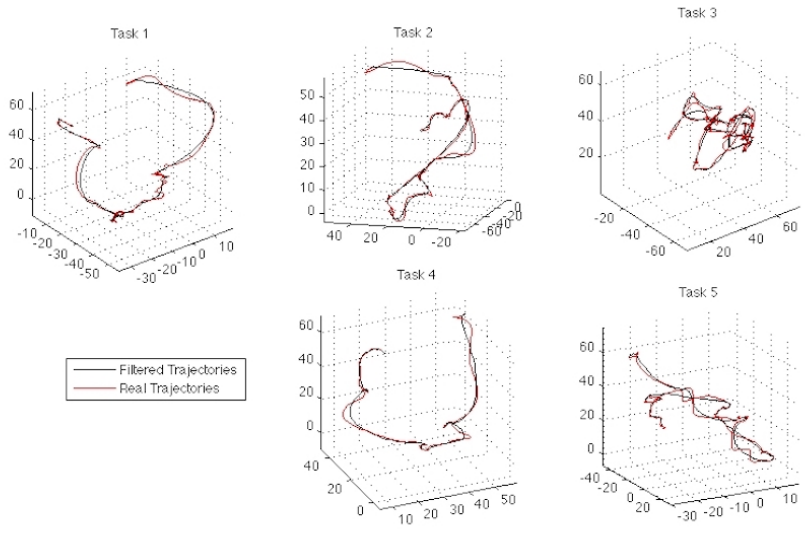

Figure 8: Different surgical gestures.

A filtering stage was applied according to the approach explained in Section 2.3. p-values for DALW and DTW distances, calculated using the com-

parison between different and same trajectories, are 0.0106 and 0.0392, respectively. The results (p-value< 0.05) show that there is a significant difference between DALW and DTW distances computed for different and same trajectories.

A further analysis includes the calculation of DTW and DALW distances based on the configuration presented in Table 1. Figure 9 shows the results associated to the mean, maximum and minimum values of DTW and DALW distance, representing the data average and range for each group of experiments.

Results obtained in this section are quite different compared to the generic gestures described in Section 4.1. In these experiments, trajectories analyzed for each task are pretty different regardless the participant who performs them. These results mean that distances associated to the experiments, performed with different gestures by the same person, are higher compared to the experiments performed with the same gesture by different people, in contrast with results presented in Figure 6. As can be seen in the results, the distance calculated using the DALW algorithm allows an easier recognition for the experiments involving different gestures than different people. Additionally, DALW distances present a lower deviation than DTW distances (See Figure 9), allowing to distinguish more clearly among the different groups of experiments.

In order to evaluate the advantages of DALW distances in these experiments, a basic perceptron multilayer neural network, described in section 4.1,

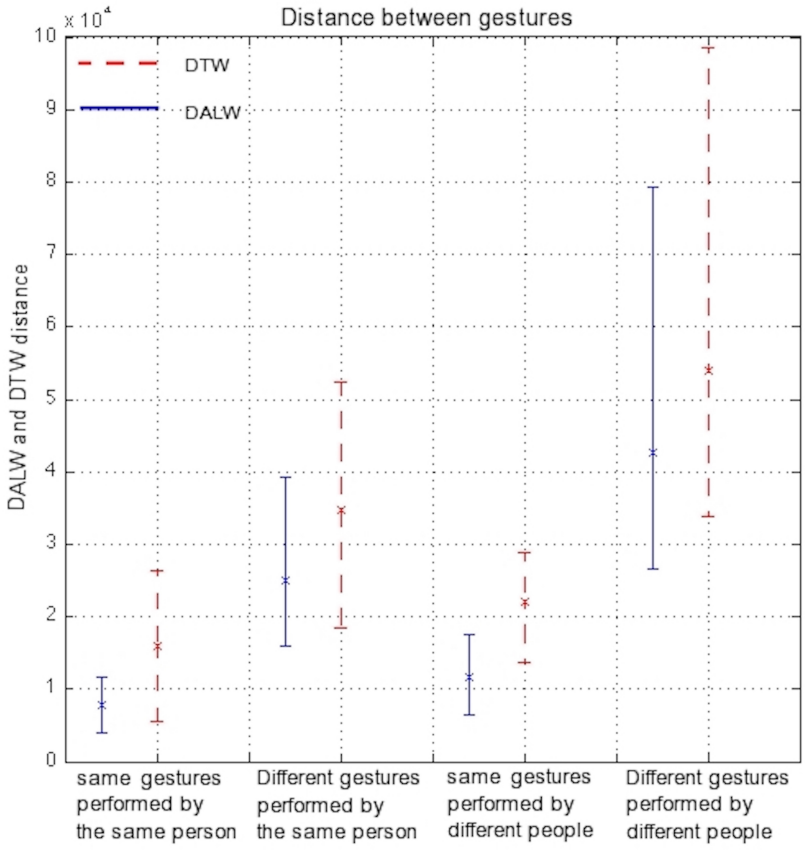

Figure 9: DTW and DALW distances calculated using position values.

was implemented. Classification rates, associated to the different 4 cases (See Table 1), were 70.1 % for DTW distances and 80.3 % for DALW distances. Althought, results implies an improvement by using the approach proposed in this work, is still difficult to distinguish among the 4 classes described previously. The reason is based on the fact that there is a big difference among trajectories based on the movement change, compared with the dynamic that could add the change of participants. Based on this idea, in order to eval-

uate the classification rates, just based on the differences associated to the movements, 2 classes were considered: different and same gestures. Results associated to DTW are described by a classification rate of 91 % and by a 98.8 %, for DALW distances.

4.3. Gesture Comparison Using an Obstetrical Forceps

This section presents an analysis of trajectories, recorded by means of an instrumented obstetrical forceps coupled with a BirthSIM simulator (See Figure 10). Using this instrument, a medical practitioner can perform a transvaginal assessment diagnosis. The BirthSIM simulator comprises anthropomorphic models of the fetal head and the maternal pelvis. Meanwhile, the forceps is instrumented to record the displacements inside the pelvis. The whole system uses electromagnetic sensors, with six Degrees of Freedom, which can track masked objects. Is important to note that forceps includes nonmagnetic material to avoid interferences in the device [20].

During these experiments, five obstetric practitioners were asked to carry out 30 forceps blade placements generating 60 trajectories, for each subject: 30 right blade and 30 left blade position signals. Trajectories were performed for 15 forceps blade placements in 2 different sessions (Figure 11). For each movement, the fetal head is located in accordance with the American College of Obstetrics and Gynecology (ACOG) classification [5] on an outlet OA+4 presentation (Occiput Anterior location and station +4cm from the ischial spines plan).

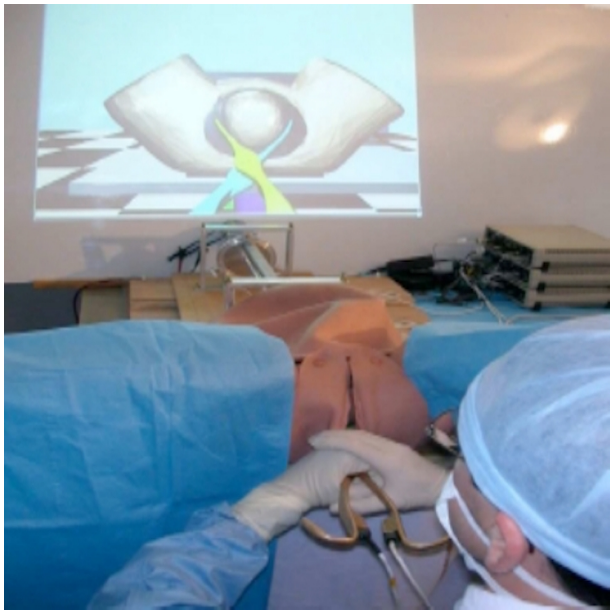

Figure 10: Obstetrical Forceps.

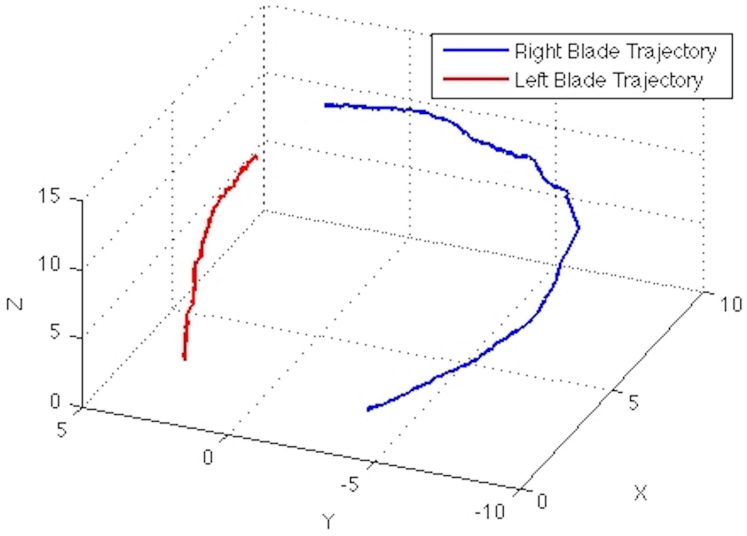

Figure 11: Left and right blade trajectories.

Once the acquisition process is completed, a low-pass filter is applied, on the basis of the approach, described in Section 2.3.p-values for DALW and DTW distances, calculated using the comparison between different and same trajectories, are 0.0061 and 0.0149, respectively. The results show that there is a significant difference between DALW and DTW distances computed for different and same trajectories (p-value<0.05).

Consequently, DALW and DTW distances were computed for each pair of trajectories, based on experiments defined in Table 1. Results, associated to DTW and DALW distances, present the mean, maximum and minimum values of the similarity measurements, for each experiment (Figure 12).

In this case, results are similar to those obtained in Section 4.2 using laparoscopic data. For the experiment performed by the same person with different gestures, distance values are higher compared to the experiment, performed by different people for the same trajectory. These results can be explained by the fact that trajectories acquired are quite different compared to the variation obtained between different participants. Consequently, 2 groups can be differentiated: those that involve different and same gestures, regardless if they were performed by different people or by the same subject. In addition, the deviation for both algorithms is smaller compared to the experiments carried out using the laparoscopic device. This result can be explained by the amount of gestures available for the analysis.

In order to evaluate the advantages of DALW distances in these experiments, a basic perceptron multilayer neural network, described in section 4.1,

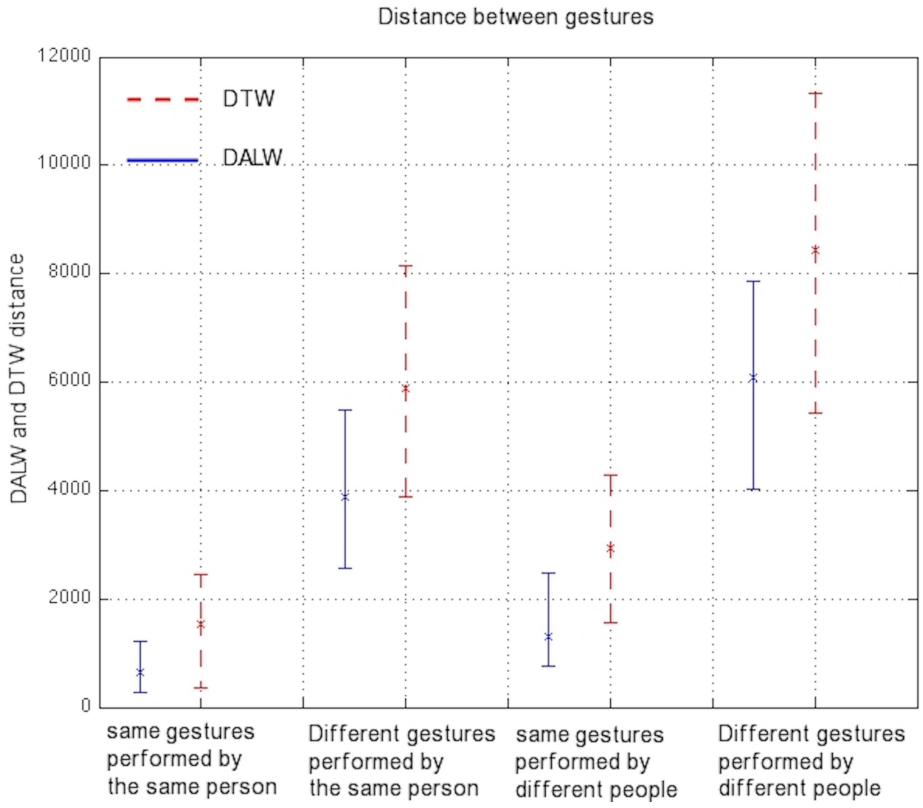

Figure 12: DTW and DALW distances calculated using position data.

was implemented. Classification rates, associated to the different 4 cases (See Table 1), were 83.5 % for DTW distances and 88.6 % for DALW distances. Althought, results implies an improvement by using the approach proposed in this work, is still difficult to distinguish among the 4 classes described previously. In this way, the classification was performed, based on those groups that can be distinguished: different and same gestures. Results associated to DTW are described by a classification rate of 96.2 % and by a 100 %, for DALW distances. Finally, as in previous experiments, DALW distances

facilitates the gestures recognition and classification for every experiment compared to DTW similarity measures.

5. Conclusion

Hand gesture recognition researches have been widely carried out and many methods are developed every year, mainly due to their applications in interactive human-machine interfaces and virtual environments in medicine.

In this work, the proposed algorithm distinguishes trajectories using a space-time independent parametrization. In particular, this method includes an arc-length parametrization providing a time independence and a geometric invariant like curvature, that can vary based on local geometry and not sensor location. This work includes trajectories carried out repeatedly by different participants; these data sets are acquired using 3 different devices. Results obtained suggest that different gestures can be distinguished based only on position measurements. Main contributions lie in the possibility to compare similar trajectories carried out at different velocities, based on the information obtained from the cumulative arc-length warping and the curvature, in order to ensure invariance to the location of the sensor coordinate system.

On the basis of these results, this new strategy can be used as an useful technique in the surgical gesture recognition and classification field. Specifically, experimental results involve an average increase of 7.14% in the classification rates by using DALW distance compared to the classical MD-DTW.

In our future work, more complex trajectories are going to be analyzed, in order to evaluate the potential of this approach, differentiating and classifying surgical gestures.

Acknowledgement

The authors would like to thank warmly Benoit Herman from Universite Catholique de Louvain, Center for Research in Mechatronics who kindly gave us some of the experimental data exploited in this work.

They wish to express their thanks, as well, to the French Ministry of Foreign and European Affairs and the National University of Colombia for the scholarships that have contributed to the development of this research

References

[1] Bankó, Z., & Abonyi, J. (2012). Correlation based dynamic time warping of multivariate time series. *Expert Systems with Applications*, *39*, 12814–12823.

[2] Berndt, D., & Clifford, J. (1994). *Using Dynamic Time Warping to Find Patterns in Time Series*. Technical Report New York University.

[3] Carmona, J. M., & Climent, J. (2012). A performance evaluation of hmm and dtw for gesture recognition. In *Iberoamerican Congress on Pattern Recognition* (pp. 236–243). Springer.

[4] Cifuentes, J., Moreau, R., Prieto, F., Pham, M., & Redarce, T. (2013). Why and how to objectively evaluate medical gestures? {*IRBM*}, *34*, 74 – 78. Digital Technologies for Healthcare.

[5] Cunningham, G., Gilstrap, L., Leveno, K., Bloom, S., Hauth, J., & Wenstrom, K. (2005). *Williams Obstetrics*. (22nd ed.). McGraw-Hill Companies. ISBN 0071413154.

[6] Datta, V., Chang, A., Mackay, S., & Darzi, A. (2002). the relationship between motion analysis and surgical technical assessment. *The American Journal of Surgery*, (pp. 70–73).

[7] Datta, V., Mackay, S., Mandalia, M., & Darza, A. (). The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *Journal of the American College of Surgeons.*, *193*, 479–485.

[8] Despinoy, F., Bouget, D., Forestier, G., Penet, C., Zemiti, N., Poignet, P., & Jannin, P. (2016). Unsupervised trajectory segmentation for surgical gesture recognition in robotic training. *IEEE Transactions on Biomedical Engineering*, *63*, 1280–1291.

[9] Frantti, T., & Kallio, S. (2004). Expert system for gesture recognition in terminal's user interface. *Expert Systems with Applications*, *26*, 189–202.

[10] Gil, J., & Keren, D. (1997). New approach to the arc length parameterization problem.

[11] Herman, B., Zahraee, A., Szewczyk, J., Morel, G., Bourdin, C., Vercher, J.-L., & Gayet, B. (2011). Ergonomic and gesture performance of robotized instruments for laparoscopic surgery. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on* (pp. 1333–1338).

[12] Holt, G. T., Reinders, M., & Hendriks, E. (2007). Multi-dimensional dynamic time warping for gesture recognition. *Thirteenth annual conference of the Advanced School for Computing and Imaging*, .

[13] King, R. C., Atallah, L., Lo, B. P., & Yang, G.-Z. (2009). Development of a wireless sensor glove for surgical skills assessment. *IEEE Transactions on Information Technology in Biomedicine*, *13*, 673–679.

[14] Kuzmanic, A., & Zanchi, V. (2007). Hand shape classification using dtw and lcss as similarity measures for vision-based gesture recognition system. In *EUROCON, 2007. The International Conference on "Computer as a Tool"* (pp. 264–269). IEEE.

[15] Levant, A. (1998). Robust exact differentiation via sliding mode technique. *Automatica*, *34*, 379–384.

[16] Levant, A. (2003). Higher-order sliding modes, differentiation and output-feedback control. *International Journal of Control*, *76*, 924–941.

[17] Li, C., Ma, Z., Yao, L., & Zhang, D. (2013). Improvements on EMG-based handwriting recognition with DTW algorithm. In *35th Annual International Conference of the IEEE EMBS* (pp. 2144–2147).

[18] Loukas, C., & Georgiou, E. (2011). Multivariate autoregressive modeling of hand kinematics for laparoscopic skills assessment of surgical trainees. *IEEE Transactions on Biomedical Engineering*, *58*, 3289–3297.

[19] Mitra, S., & Acharya, T. (2007). Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *37*, 311–324.

[20] Moreau, R., Pham, M. T., Silveira, R., Redarce, T., Brun, X., & Dupuis, O. (2007). Design of a new instrumented forceps: Application to safe obstetrical forceps blade placement. *IEEE Transactions on Biomedical Engineering*, *54*, 1280–1290.

[21] Moubarak, S., Pham, M., Moreau, R., & Redarce, T. (2010). Gravity compensation of an upper extremity exoskeleton. In *In International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 10)* (pp. 4489–4493). Buenos Aires, Argentina.

[22] Moubarak, S., Pham, M., Pajdla, T., & Redarce, T. (2009). Design and modeling of an upper extremity exoskeleton. In *In 11th International Congress of the IUPESM: Medical Physics and Biomedical Engineering World Congress 2009* (pp. 476–479). Munich, Germany.

[23] Mueen, A., & Keogh, E. (2016). Extracting optimal performance from dynamic time warping. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 2129–2130). ACM.

[24] Murphy, T. E. (2004). *Towards objective surgical skill evaluation with hidden Markov model-based motion recognition*. Ph.D. thesis Johns Hopkins University.

[25] Pan, L., Sheng, X., Zhang, D., & Zhu, X. (2013). A linear model for simultaneously and proportionally estimating wrist kinematics from EMG during mirrored bilateral movements. In *35th Annual International Conference of the IEEE EMBS* (pp. 4593–4596).

[26] Pham, M., Moreau, R., & Boulanger, P. (2010). Three-dimensional gesture comparison using curvature analysis of position and orientation. *In IEEE-EMBS International Conference of the Engineering in Medicine and Biology Society.*, (pp. 6345–6348).

[27] Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, *43*, 1–54.

[28] Reyes, M., Dominguez, G., & Escalera, S. (2011). Featureweighting in dynamic timewarping for gesture recognition in depth data. In *Com-*

puter Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on (pp. 1182–1188). IEEE.

[29] Richards, C., Rosen, J., Hannaford, B., Pellegrini, C., & Sinanan, M. (2000). Skills evaluation in minimally invasive surgery using force/torque signatures. *Surgical Endoscopy.*, *14*, 791–798.

[30] Rosen, J., Hannaford, B., Richards, C., & Sinanan, M. (2001). Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. *IEEE Transactions on Biomedical Engineering*, *48*, 579–591.

[31] Saggio, G., Santosuosso, G., Cavallo, P., Lorenzo, N., Lazzaro, A., & Corona, A. (2011). Gesture recognition and classification for surgical skill assessment. *IEEE International Workshop on Medical Measurements and Applications Proceedings (MeMeA)*, (pp. 662–668).

[32] Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, *26*, 43–49.

[33] Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *27*, 43–49.

[34] Singh, B., & Singh, H. K. (2010). Web data mining research: a survey.

In *Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on* (pp. 1–10). IEEE.

[35] Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. *Communications of the ACM*, *54*, 60–71.

[36] Wang, N., Chen, Y., & Zhang, X. (2013). The recognition of multi-finger prehensile postures using lda. *Biomedical Signal Processing and Control*, *8*, 706–712.

[37] Woolson, R. (2008). Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials*, .

[38] Yang, D., Zhao, J., Gu, Y., Jiang, L., & Liu, H. (2009). Emg pattern recognition and grasping force estimation: Improvement to the myocontrol of multi-dof prosthetic hands. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on* (pp. 516–521). IEEE.