An Arc-length Warping Algorithm for Gesture Recognition Using Quaternion Representation

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

Abstract— **This paper presents a new algorithm, called Dynamic Arc-Length Warping algorithm (DALW) for hand gesture recognition based on the orientation data. In this algorithm, after calculating the quaternion for each orientation measurement, we use DALW algorithm to obtain a similarity measure between different trajectories. We present the benefits of using quaternion alongside the implementation of Dynamic Arc Length Warping to present an optimized tool for gesture recognition. We show the advantages of this approach compared with other techniques. This tool can be used to distinguish similar and different gestures. An experimental validation is carried out to classify a series of simple human gestures.**

I. INTRODUCTION

Hand posture and gesture recognition is an essential mechanism for modern human-machine interaction. The range of applications is quite broad. From the identification of sign language to medical rehabilitation many researchers [1], [2], [3] have developed new algorithms and systems to obtain a natural interaction between humans and computers. In order to do so two tools are required: an interface for the acquisition of the movements and an accurate hand gesture recognition system. The recognized gestures can be used for controlling a robot or conveying meaningful information [4] to the computer.

Data acquisition can be performed by either using imaging or by using specific devices such as instrumented exoskeleton [5]. One advantage of specific devices is that they are not affected by the surroundings environment but tends to be quite intrusive. However, both approaches are widely used to extract gesture features [6] that can then be recognized. The data sets acquired with these systems are interpreted and analyzed using different algorithms in order to perform an adequate classification and comparison. To detect static gestures a general classifier or a template-matcher can be used. However, dynamic gestures classification requires tools that can handle multidimensional temporal data [7], [8]. The methodology developed in [9] uses a subset of hidden Markov modeling to perform this task. In this work, force and torque measurements are used to define a force/torque signature associated with 14 different gestures in surgery. This approach has also been used to recognize a forty word subset of the American Sign Language [10]. Other statistical approaches based on Dynamic Time Warping (DTW) [11] or Longest Common Subsequence (LCSS) [12] have been used to find the optimal alignment of two temporal signals. Several advances have been made in the surgery field with these tools. Pham et al. [13] use DTW to match and compare the curvature of two 3-D trajectories successfully.

In human gesture recognition, full hand localization can be achieved not only with finger's position information but also with their orientation information. Particularly, for the evaluation of surgical gestures, the orientation of the tip of the instrument is essential to determine how tissues are manipulated in a surgical procedure. DTW and its variations were successfully applied to position data [8] but not to orientation data, which is the purpose of this paper. In this paper, we use a quaternion representation of the angles to classify a gesture. The main advantages are related to its coordinate system independence and the fact that the interpolation between two quaternions can be performed easily [17]. To perform a comparison between different sets of data, we use a new warping algorithm called Dynamic Arc-Length Warping (DALW). This algorithm starts by an arc-length reparametrization allowing time independence and then uses a geometric invariant like curvature that only varies according to local geometry and not sensor location. Moreover, this technique allows the dimension of the problem to be reduced from 3D to 1D data simplifying the classification problem. Good results have been obtained with this algorithm in comparison with the classical Dynamic Time Warping with position datasets [15]. In this work, we take advantage of the benefits of Dynamic Arc-Length Warping and the use of quaternions to obtain an optimal tool for classification of gestures based on their values orientation. It is important to note that this approach has been applied to a set of data recorded offline. If the application implies an online configuration, advanced robust techniques of differentiation are necessary. Similarly, if the data contains significant noise, it is necessary to implement a robust filtering stage in order to obtain smooth trajectories to test the algorithm.

The following section presents the mathematical foundation of orientation data and quaternion. Section 3 describes Dynamic Arc-Length algorithm. Section 4 presents experimental results and an analysis of the results. In the last section, we conclude and propose future work.

II. USING QUATERNION TO REPRESENT ORIENTATION

Multi-dimensional temporal series are data sets with multiple measurements made simultaneously that varies over time. They are typically a vector of feature values for each time occurrence. In this paper, we carry out an analysis of

[1]Laboratory Ampere, UMR CNRS 5005, Institut National des Sciences Appliquées de Lyon, France

[2]Department of Mechanic and Mechatronic Engineering, Universidad Nacional de Colombia, Colombia.

[3]University of Alberta, Canada

orientation data of the wrist produced by an exoskeleton measuring exactly the orientation information in real-time.

Orientation of an object in three-dimensional space can be parameterized using quaternion. A quaternion **Q** is defined by [16]:

$$Q = a + i.b + j.c + k.d \tag{1}$$

where $a, b, c, d \in R$ and i, j, k are different imaginary units. A rotation can be described based in the rotation axis L, its unit vector $l = [l_x, l_y, l_z]^T$ and the rotation angle ϕ with the unit quaternion as described by:

$$Q = \cos\frac{\phi}{2} + i.l_x\sin\frac{\phi}{2} + j.l_y\sin\frac{\phi}{2} + k.l_z\sin\frac{\phi}{2}. \tag{2}$$

The main advantage of quaternion is that its representation is independent of a central coordinate system which avoids singular situations such as the so called *gimbal lock* for Euler angles. A quaternion is also more compact and faster than the traditional matrix representations.

One can calculate a quaternion based on orientation measurements using the following expression [14]:

$$Q = \begin{bmatrix} \cos(\alpha/2)\cos(\beta/2)\cos(\gamma/2) + \sin(\alpha/2)\sin(\beta/2)\sin(\gamma/2) \\ \sin(\alpha/2)\cos(\beta/2)\cos(\gamma/2) - \cos(\alpha/2)\sin(\beta/2)\sin(\gamma/2) \\ \cos(\alpha/2)\sin(\beta/2)\cos(\gamma/2) + \sin(\alpha/2)\cos(\beta/2)\sin(\gamma/2) \\ \cos(\alpha/2)\cos(\beta/2)\sin(\gamma/2) - \sin(\alpha/2)\sin(\beta/2)\sin(\gamma/2) \end{bmatrix} \tag{3}$$

where α, β and γ are the angles around the axes x, y and z.

III. Dynamic Arc-Length Warping

With the data described previously, a new algorithm called Dynamic Arc-Length algorithm is implemented to perform gesture classification.

A. Arc-length Parametrization

In the proposed approach, first an arc-length parametrization is performed.The new parametrization is quite useful because the cumulative arc length arises naturally from the shape of the curve and does not depend on a particular coordinate system. For the orientation trajectories represented by the vector $\mathbf{Q}(t) = [\alpha(t), \beta(t), \gamma(t)]$, the cumulative arc-length is based on the total length L of the orientation trajectory between the beginning and the end of the gesture $[t_a, t_b]$ and is defined by:

$$S = \int_{t_a}^{t_b} ||\mathbf{Q}'(t)||dt \tag{4}$$

where $\mathbf{Q}'(t)$ is the first derivative of the orientation trajectory with respect to time.

The trajectory can be re-parameterized using a normalized parameter s called the cumulative arc-length defined by:

$$s = \frac{1}{S}\int_{t_a}^{t_b} ||\mathbf{Q}'(t)||dt \tag{5}$$

with $s \in [0, 1]$.

For the quaternion, we calculate the distance between two successive quaternions to compute the corresponding arc-length. In this case, we compute the unit quaternion that satisfies:

$$a^2 + b^2 + c^2 + d^2 = 1 \tag{6}$$

and the distance between two quaternions Q_1 and Q_2 is computed as:

$$d(Q_1, Q_2) = 1 - \langle Q_1, Q_2\rangle^2 \tag{7}$$

where $\langle Q_1, Q_2\rangle$ denotes the inner product between the two quaternions.

B. Quaternion Curvature

After computing the arc-length parametrization, we reduce the problem dimensionality by obtaining the curvature from each data set. The curvature vector $\mathbf{Q}''(s)$ is defined as the rate of change of the unit tangent vector with respect to the arc-length. We define curvature as the length of the curvature vector. We will denote the curvature scalar quantity by the letter κ as following:

$$\kappa(s) = ||\mathbf{Q}''(s)||. \tag{8}$$

When the parametrization with respect to the arc-length is used, the trajectory is sampled with unequal intervals in space. For this reason, one cannot apply the classical techniques of numerical differentiation to compute the curvature. One way to handle the problem is to fit the data using polynomial functions. In our case, a second-order Lagrange polynomial interpolation scheme is used to fit each set of three adjacent points using:

$$f_n(s) = L_{i-1}(s)p_n(s_{i-1}) + L_i(s)p_n(s_i) + L_{i+1}(s)p_n(s_{i+1}). \tag{9}$$

Where the function f_n is the Lagrange polynomial, L_i is the Lagrange basis function, p_n are the points to interpolate and s_{i-1}, s_i, s_{i+1} is the cumulative arc-length of three consecutive points.

Using this numerical approximation, one can calculate the numerical derivatives of two trajectories $\mathbf{Q}_1(s')$ and $\mathbf{Q}_2(s')$ in order to obtain their corresponding curvatures.

C. Dynamic Arc-Length Warping

The main objective of Dynamic Arc-Length Warping **DALW** is to compare two curvature sequences as expressed by:

$$\kappa_1 = (\kappa_1(s_1), \kappa_1(s_2), \kappa_1(s_3), ..., \kappa_1(s_l))$$

$$\kappa_2 = (\kappa_2(s'_1), \kappa_2(s'_2), \kappa_2(s'_3), ..., \kappa_2(s'_m))$$

where the lengths $l \in \mathbb{N}$ and $m \in \mathbb{N}$, respectively.

To align two different curvature signatures κ_1 and κ_2, a local comparison 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 work, we will use the Euclidean norm for convenience. This local similarity distance is defined between any pair of elements $\kappa_1(s_i)$ and $\kappa_2(s'_j)$ using the following expression:

$$F(i,j) = (\kappa_1(s_i) - \kappa_2(s'_j))^2 \geq 0 \tag{10}$$

Typically $F(i,j)$ is small (low cost) if $\kappa_1(s_i)$ and $\kappa_2(s'_j)$ are similar to each other and $F(i,j)$ is large otherwise(high cost). While evaluating the local cost measure between two sequences κ_1 and κ_2 leads to define a cost matrix by $\mathbf{F}$. Then main goal is to find an alignment between κ_1 and κ_2 which minimize the overall cost function. Let us define a warping path of a curvature signature κ_n by $\phi_n(k),\ k = 1....W$ which is defined by:

$$\kappa'_n(s_j) = \kappa_n(s_{\phi(k)}) \tag{11}$$

where κ'_n is a new curvature array where the s_j value corresponds to the previous signature value at s_k.

Using this notation the warping functions $\phi_1(k)$ and $\phi_2(k)$ provide a new mapping of the cumulative arc-length index of κ_1 and κ_2 respectively. Given a mapping ϕ, one can compute the total distance between the warped arc-length series as:

$$d(k) = \sum_{k=1}^{W} \frac{1}{\alpha_{\phi(k)}} (\kappa_1(s_{\phi_1(k)}) - \kappa_2(s_{\phi_2(k)}))^2 M_\phi \tag{12}$$

where $\alpha_{\phi(k)}$ is a per-step weighting coefficient and M_ϕ is the corresponding normalization constant, which ensures that the distances are comparable along different paths.

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

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

This result is equivalent to the distance or similarity measure between paths. This optimization problem can be efficiently solved by using a dynamic programming technique.

IV. EXPERIMENTAL RESULTS

To test our method, simple gestures were studied. The orientation measurement for each gesture were recorded using an exoskeleton device described in [5]. Figure 1 shows the exoskeleton used and its kinematic model. This device has four degrees of freedom: three for the shoulder (internal-external rotation, abduction-adduction, flexion-extension) and one for the elbow (flexion-extension).

A total of 12 subjects were involved in this study. They were asked to do two kinds of gesture (10 times each) with their right arm (all subjects are right handed). The first one was a communication gesture as the subject wanted to say stop to someone running towards him. The second one was a simple gesture where the subject had to place their hand on a wall in front of them. Figure 2 represents an example of the average values of orientation trajectories of a subject's wrist.

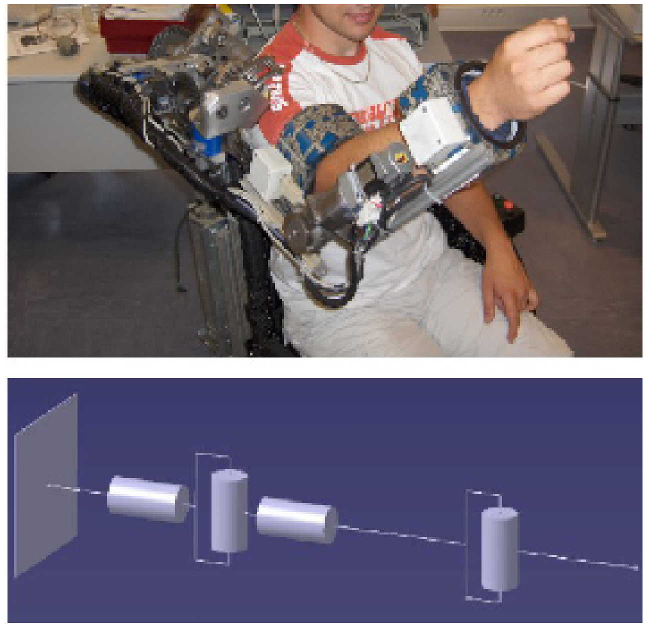

Fig. 1. The upper extremity exoskeleton with its four active degrees of freedom.

The angular orientation trajectories and the corresponding quaternion were aligned using the DALW algorithm according to four different modalities (Table I).

TABLE I

DIFFERENT EXPERIMENTS

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

The DALW distance is computed with orientation data and with the corresponding quaternion according to these four criteria.

The results of DALW distance values and the standard deviations for each experiment can be seen in Figure 3.

As one can see in Figure 3, the DALW distance values obtained with quaternion differentiate clearly the different sets of experiments allowing a classification of the gestures. Clearly, the distance measured is higher for the experiment which involves different gestures performed by different people and lower for the experiment which compares similar gestures performed by the same person. Likewise, this similarity measure allows to differentiate the experiments involving different gestures performed by the same person or by different people. Also, it is clear that the DALW algorithm with quaternion leads to smaller standard deviations compared to the DALW algorithm with the angular orientation data allowing obtain data sets more clustered that facilitate the classification stage.

The results obtained in this work allow to take into account the quaternion information from angular orientation

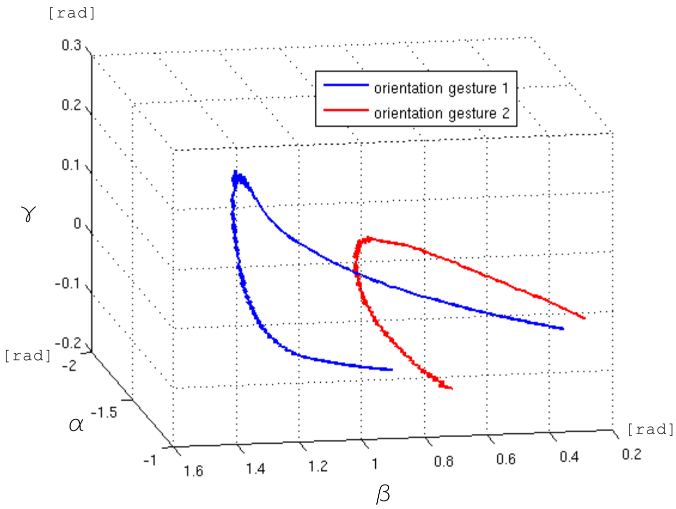

Fig. 2. An example of orientations of the wrist.

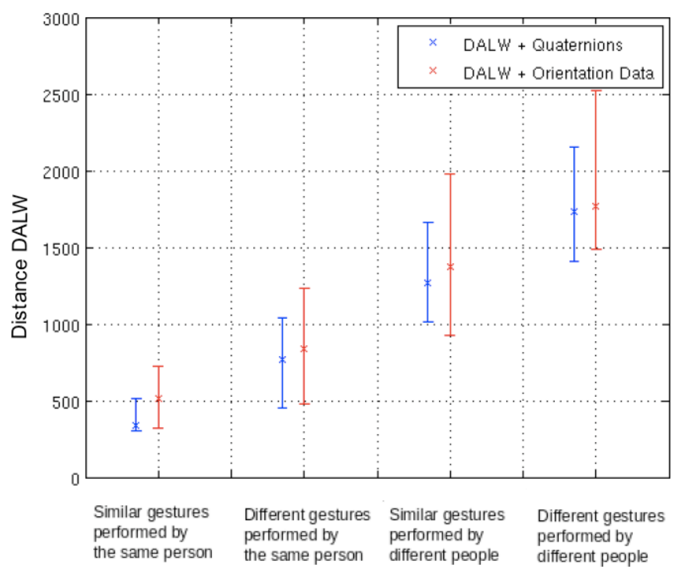

Fig. 3. DALW Distance found with quaternions and orientation data

data and the Dynamic Arc-Length Warping algorithm as an appropriate tool for the hand human gesture recognition.

V. Conclusion

In this paper, we have presented a hand gesture recognition method that combines a quaternion representation and a Dynamic Arc-Length Warping algorithm to perform gesture classification. The use of quaternion and arc-length parametrization is invariant to the sensor coordinate system. Additionally, our approach performs a dimensionality reduction making the recognition task simpler.

We have also presented an experimental study with two different gestures to validate our approach. A total of 12 subjects were involved with 10 attempts for each gesture. The results show that distances obtained with a quaternion representation lead to a standard deviation lower than those obtained with angular orientation data. Furthermore, the higher standard deviations while using the angular orientation data are greater for the comparison between different groups of experiments.

The similarity measure obtained between the different sets of data allows to develop a relevant classification of human gestures. Overall, the results highlight this approach as a good tool for gesture recognition as well.

In the near future, we will continue to work on increasing the number of hand gestures recognized by the system. Also, our ongoing research is focusing on the study of a considerable amount of real surgical gestures measured by instrumented endoscopes in minimally invasive surgery procedures. We hope that this technique will allow us to improve the objective assessment of surgical gestures.

References

[1] R. Bowden, D. Windridge, A. Zisserman and M. Brady, linguistic feature vector for the visual interpretation of sign language, European Conference on Computer Vision, Vol. 1, pp 391- 401, 2004.

[2] T. Baudel M. Lafon, Charade: Remote control of objects using free-hand gestures, Communications of the ACM, Vol. 36, pp 28-35, 1993.

[3] I. Allen, Doctors and their careers, Policy Studies Institute, 1994.

[4] G. R. S. Murthy and R. S. Jadon, A Review of Vision Based Hand Gestures Recognition, International Journal of Information Technology and Knowledge Management, Vol. 2, No. 2, pp 405- 410, 2009

[5] S. Moubarak, M. Pham, R. Moreau and T. Redarce, Gravity compensation of an upper extremity exoskeleton, In In International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 10) (p. 44894493), Buenos Aires, Argentina

[6] G. Wysoski, V. Lamar, S. Kuroyanagi, and A. Iwata, A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural Networks, IEEE 9th International Conference on Neural Information Processing, Singapur, 2002.

[7] A. Corradini, Dynamic time warping for offline recognition of a small gesture vocabulary, In Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, July-August, 2001.

[8] J. Cifuentes, M.Pham, F. Prieto, P. Boulanger and R. Moreau, Objective Assessment of Surgical Skills, 2012 Proceedings of the ASME 11th Biennial Conference on Engineering Systems Design and Analysis (ESDA2012), July, Nantes, France, 2012.

[9] J. Rosen,B. Hannaford, C. Richards, and M. Sinanan, A Rotation Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills, IEEE Transactions on Biomedical Engineering, Vol. 48, pp 579-591, 2001.

[10] T. Starner, Visual recognition of american sign language using hidden markov models, Master's Thesis, Massachusetts Institute of Technology, 1995.

[11] J. Kruskall and M. Liverman,The symmetric time warping problem: From continuous to discrete, In Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison, pp. 125-161, 1983.

[12] L. Bergroth, H. Hakonen and T. Raita, A survey of longest common subsequence algorithms, Seventh International Symposium on String Processing and Information Retrieval, pp. 39-48, 2000.

[13] M. Pham, R. Moreau and P. Boulanger, 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, 2010

[14] G. Gubaidullin, Euler angles and quaternions in Robotics, Department of Mathematics and Technology, University of Applied Sciences Koblenz, Germany, 2004

[15] J. Cifuentes, R. Moreau, F. Prieto, M.T. Pham and T. Redarce, Why and how to objectively evaluate medical gestures?, IRMB, Vol. 34-1, pp 74-78, February 2013

[16] J. Kuipers. Quaternions and Rotation Sequences, Coral Press, pp 127-143, September 1999.

[17] A. Hanson. Visualizing Quaternions, Morgan Kaufmann Publishers, Elsevier, 2006.