# Approximation of curvature and velocity using adaptive sampling representations - Application to hand gesture analysis

Sylvie Gibet, Pierre-François Marteau

VALORIA, Université de Bretagne Sud, Campus de Tohannic, rue Yves Mainguy, F-56000 Vannes, France
{Sylvie.Gibet, Pierre-Francois.Marteau}@univ-ubs.fr

**Abstract.** This paper describes a new approach to analyze hand gestures, based on an experimental approximation of the shape and kinematics of compressed arm trajectories. The motivation of such model is on the one hand the reduction of the gesture data, and on the other hand the possibility to segment gestures into meaningful units, yielding to an analysis tool for gesture coding and synthesis. We show that the measures of the distance between adaptive samples and velocity estimated at these points are respectively correlated to the instantaneous curvature and tangential velocity directly computed on motion capture data. Based on these correlation results, we propose a way to identify an appropriate compression rate of the adaptive sampling algorithm. We also show that this new analysis tool can be applied on multidimensional data.

## 1 Introduction

The representation and the accurate understanding of human gesture is a crucial and challenging problem which was raised in several research fields, including animation of embodied agents, sport sciences, medicine or vision-based recognition. In recent years, the huge development of new technologies for motion capture has made the analysis of human motion feasible, and yielded to data-based methods for gesture classification, retrieval, and computer-generated animation.

One major problem in representing gesture from recorded data is that these data are multidimensional and direct use of them is rather expensive and fastidious. Another problem is the lack of flexibility. Computing new motion from real motion clip necessitates the elaboration of huge data sets, or the development of data-based methods for editing, blending or adapting existing motion. Finally, finding the best motion representation is a central problem, depending on the application. In particular, extracting regular features, identifying segmental units or defining measures for comparing two similar motion clips can be useful for retrieval or recognition process. As these processes operate on multidimensional data, one way to characterize gesture is to compress the original information, and to use this data reduction to characterize significant movement units. The automatic extraction of targets [1] is also an efficient way to synthesize new gestures, which takes into account the spatial variability of gestures and the co-articulation effects.

Motion capture data generally consist of sampled trajectories for each degree-of-freedom characterizing the position and orientation of the human joints. These joint data can be represented by different sets of coordinates; in particular, angular coordinates generally expressed by Euler angles or quaternion, and Cartesian coordinates.

In this paper, as we are mainly interested in visual gestures, which are gestures that "draw" the 3D space, we express them by 3D Cartesian trajectories. These gestures are most of the time conveying meaningful information, as in sign languages gestures, or expressive gestures like dance or musical gestures. They can be characterized by their shape (change of curvature), as well as by their kinematic specificities. In sign language gestures for example, the signer can draw the shape of the symbol as an icon of some aspect of the object or the activity to be symbolized (Fig. 1.a and b). Expressive gestures may also implicitly contain some velocity or acceleration profiles. In particular variations in velocity are responsible for the aggregation of samples in some areas of the trajectories (Fig. 1.c).

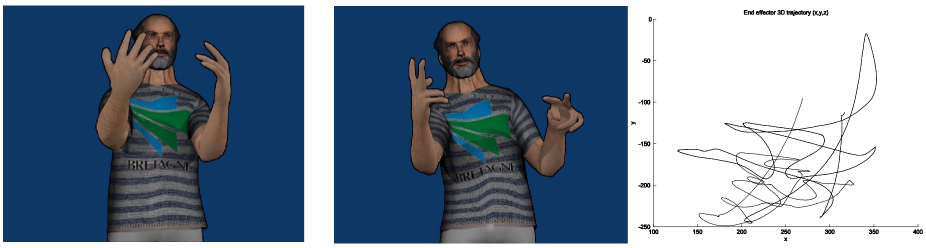



**Fig. 1.** left and middle: gesture "drawing" the space ; right: 3D end-point trajectory

We propose here to study both these spatial and kinematics characteristics in a reduced representation space. First of all we will consider the arm end-point trajectory, and the method can then be extended to multidimensional arm trajectories. While basing our work on a compressed representation of trajectories, we define an approximation of adaptive velocity and curvature. We show that these approximations can be strongly related to curvature and tangential velocity, not only in 3D space, but in any dimensional space. These measures provide new tools to automatically analyse gestures. An interpretation is given for the segmentation of sign language gestures and it possible use for gesture representation and synthesis.

The paper is mainly composed of four sections. Section 2 gives an overview of the related works. After presenting briefly the adaptive non uniform sampling algorithm used for data reduction, section three presents the analysis method. Section 4 presents some analysis for 3D arm end-point trajectories, in terms of correlation and compression rate. After illustrating some results about the segmentation of sign language end-point trajectories in section five, the possibility to use this segmentation for different purposes is evoked. The paper concludes and gives some perspectives in section six.

## 2 Related works

Numerous techniques have been developed for the analysis of human motion capture data. These studies differ considerably, whether the emphasis is placed on the search of regular features for explaining neural mechanisms, or data reduction, or on gesture segmentation for animation purposes.

The search for invariant features has been largely investigated in recent years. Researchers tried to express these regularities in terms of motion laws. Some of them can be used for trajectory segmentation. In particular, the two-third power law, expressing a power relation between velocity and curvature [2] was proposed for segmenting three dimensional unconstrained drawing movements, on the basis of abrupt changes of the velocity gain factor. Another segmentation hypothesis was based on the observation that endpoint trajectories of human arm movements tend to be piecewise planar [3]. These segmentation hypotheses are largely discussed in the neuroscience community.

There are many different mathematical approaches for curves and surfaces approximations, which tend to reduce the dimensionality of the motion data. Few works concern motion trajectories. Polygonal approximation provides characteristics points to represent the shape of the trajectory. These points, which correspond to local curvature extrema, can be connected by line segments. This method has been used by [4] for non-uniform sub-sampling of motion time-series. Another method proposes curve approximation using active contours [5]. These methods are developed for dance gesture recognition.

Other methods have been proposed to the problem of approximating multidimensional curves using piecewise linear simplification and dynamic programming in $O(kn^2)$ complexity [6]. Some efficient algorithms [7-8] (in $O(nlog(n))$ complexity) have been proposed.

Independently of the data reduction method, we propose in this paper to characterize gesture trajectories expressed in the reduced space by approximated measures of spatial shape and kinematics.

## 3 Analysis of arm movements

The gestures consist of raw data composed of 3D Cartesian trajectories, each trajectory representing the evolution with time of one coordinate *x*, *y*, or *z* expressing the position of a specific joint. For our study, we consider $X(t)$ as constituted of time-series in 3.*p* dimensions, represented by spatial vectors $X(t) =[x_1(t), y_1(t), z_1(t)\ x_2(t), y_2(t), z_2(t)... x_p(t), y_p(t), z_p(t)]$. In practice, we deal with the sampled trajectory at a constant frequency of 120 Hz: $X(n)$ where $n$ is the time-stamp index.

In this analysis method, we rely on the algorithm described in [1], which is related to linear piecewise curve approximation. This algorithm finds samples in the time series $X(t)$, not regularly located in time, as shown in figure 2.

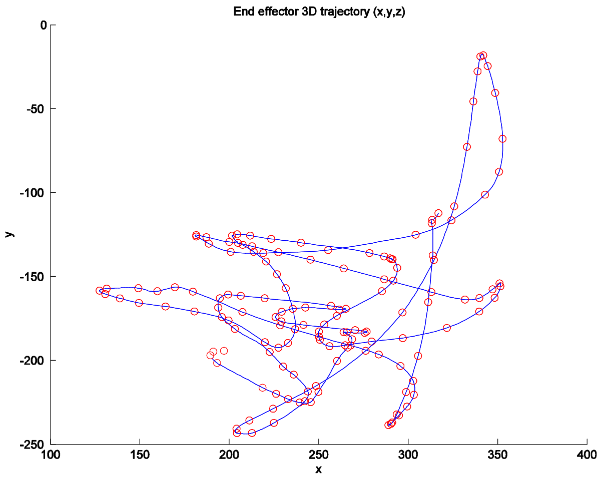



**Fig. 2.** An approximation of the end-point trajectory using the non uniform sampling algorithm

The approach consists in seeking an approximation $X_{\hat{\theta}}$ of *X(n)*, $\theta$ being the set of discrete time location $\{n_i\}$ of the segments' endpoints. The selection of the optimal set of parameters $\hat{\theta} = \{\hat{n}_i\}$ is performed using a dynamic programming algorithm. The result of this method is the optimal identification of discrete $X_{Ti}$ key-points – we call them spatial targets – delimitating the segments, for a given compression rate $\rho$. The complexity of the algorithm is $O(n^2/k)$ where $n$ is the number of samples, and $k$ the number of segments, but can be decreased down to $O(n)$ if optimality is somehow relaxed [9]. For detailed description of the method, see [1].

In a first study we work on 3D end point trajectories $X(t) = [x(t), y(t), z(t)]$, the coordinates being calculated in the shoulder frame. For any smooth trajectory parameterized with $t$, we express the instantaneous velocity $v(t)$ and the absolute value of the instantaneous curvature $\kappa(t)$:

$$v(t) = \left\| \dot{X}(t) \right\| = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \quad (1)$$

$$\kappa = \frac{\left\| \dot{X}(t) \times \ddot{X}(t) \right\|}{\left\| \dot{X}(t) \right\|^3} \quad \text{and} \quad R(t) = \frac{1}{|\kappa|} \qquad (2)$$

where $R$ is the radius of curvature. The curvature measures how fast a curve is changing direction at a given point.

These variables have been extensively studied for a variety of goal-directed experimental tasks. In particular, a number of regularities have been empirically observed for end-point trajectories of the human upper-limb, during 2D drawing movements.

However, for 3D movements with great spatial and temporal variations, it can be difficult to directly extract significant features from these signals. Moreover, computing the radius of curvature raises a problem, when the velocity is too high, or when there are inflexion points in the trajectories. In particular for noisy data the radius of curvature may be difficult to compute. Finally, for higher dimensions, the curvature is not defined, prohibiting its use in the angular space in particular.

We propose to approximate these velocity and curvature by empirical measures calculated from the adaptive samples identified through the *DPPLA* algorithm [1]. We define the target-based velocity by the expression:

$$V_{T_{gi}}(n_i) = \frac{\left\| X(n_{i+1}) - X(n_{i-1}) \right\|}{n_{i+1} - n_{i-1}} \qquad (3)$$

where $n_{i+1}$ and $n_{i-1}$ are temporal indices of the associated targets $Tg_{i+1}$ and $Tg_{i-1}$. As the targets are not regularly located, the addition effect of this measure, homogeneous to a velocity, is to filter the raw data. The filtering depends on the compression rate.

We define as well the inverse distance between adjacent targets as:

$$\kappa_{T_{gi}}(n_i) = \frac{1}{\left\| X(n_i) - X(n_{i-1}) \right\|} \qquad (4)$$

With this formulation, we assume that this last quantity might be linked to a measure of aggregation points on the trajectory: when the movement velocity decreases, the distance between original samples decreases and the curvature appears to be important. Therefore, $\kappa_{Tgi}(n_i)$ expresses a spatial quantity which might be correlated to curvature at time-index $n_i$.

In the next section, we will study the correlation between the target-based approximations and the instantaneous values. We will also study the influence of the compression parameter $k$ of the compression algorithm.

# 4 Analysis of 3D endpoint arm data

*Corpus*

One deaf signer performed the gestures. He signed sequences of French sign language gestures representing several versions of bulletin weather performed with different styles, relative to the subject's dynamics and emotional state. The sequences were composed of 12 phrases; the whole duration was about 30 s. The subject was asked to perform the gestures with variations of the geometry (large vs. small amplitude), kinematics (high vs. low speed) and dynamics (smooth vs. jerky).

*Pre-processed data*

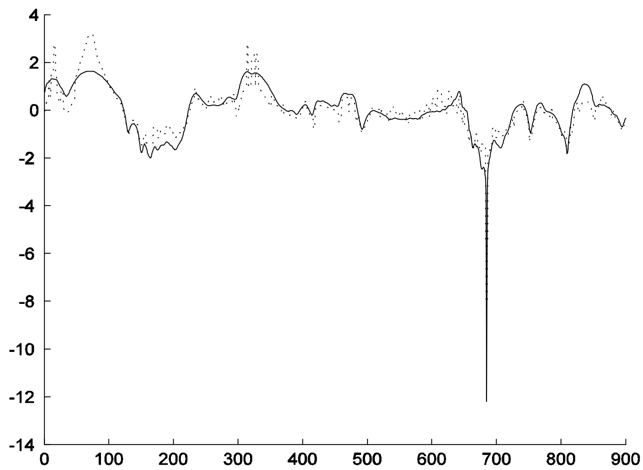
Raw data are first filtered by a low pass Butterworth filter with a cutoff frequency of 10.0 Hz. We consider sequences of about 10000 frames.

*Correlation between approximated and instantaneous variables*

The analysis of correlation is achieved, on the one hand between the log of target-based velocity and the log of its instantaneous value, and on the other hand between the inverse of the distance between targets and the instantaneous curvature. The results concerning the velocity are shown in figure 3 (a). They illustrate an excellent correlation between the two variables, thus allowing us to use target-based velocity as a good approximation of instantaneous velocity. We may also compute the acceleration of arm end-point trajectories on the basis of this target-based velocity.

The correlation between the log of the inverse target distances and the log of its instantaneous curvature is also very good, as illustrated in figure 3 (b). The points with abrupt changes are located at the same place, but the target-based signal seems less noisy than the original one. This makes possible to approximate curvature as the inverse of target density.

a) Tangential velocity (solid line) vs. target-based density (dotted line)




b) Curvature (solid line) vs. target density (dotted line)

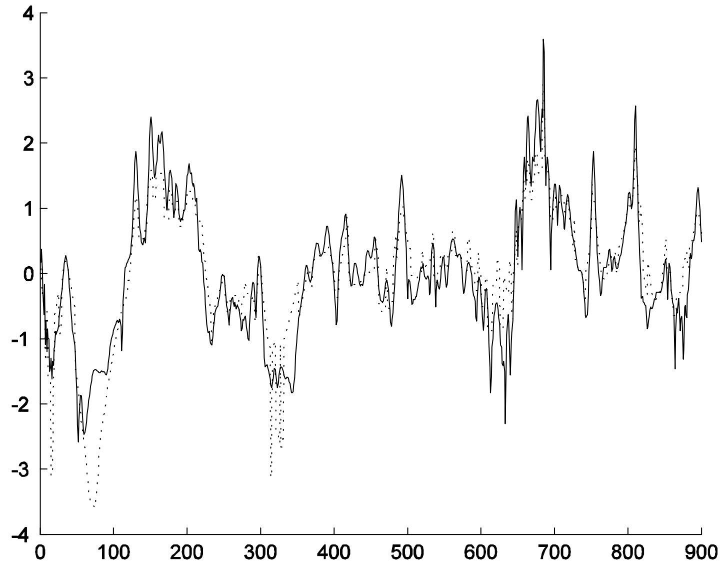



**Fig. 3.** Correlation for 3D end-point trajectories of arm movements; a) correlation between instantaneous tangential velocity (solid line) and target-based velocity (dotted line); b) correlation between instantaneous curvature (solid line) and inverse target density (dotted line)
For each signal *x*, we computed: (log(*x*) – mean(log(*x*)))/std(log(*x*))

*Influence of the compression coefficient on the correlations*

The influence of the compression factor characterizing the adaptive sampling algorithm is analyzed at the light of the correlation coefficient. The results can be seen in figure 4. It shows that for the target-based velocity, the correlation coefficient remains very close to 1, independently of the compression rate (from 50% to 95%). For the target-based acceleration, the correlation coefficient is very good (0.9), for a compression rate varying until 70%. Beyond this limit, the correlation coefficient abruptly falls. The correlation coefficient is lower for the inverse distance, but still high (.85), even for a high compression rate (until 80%). These results support the assumption that target-based variables can be used without a significant loss of data for the analysis of 3D end-point trajectories.

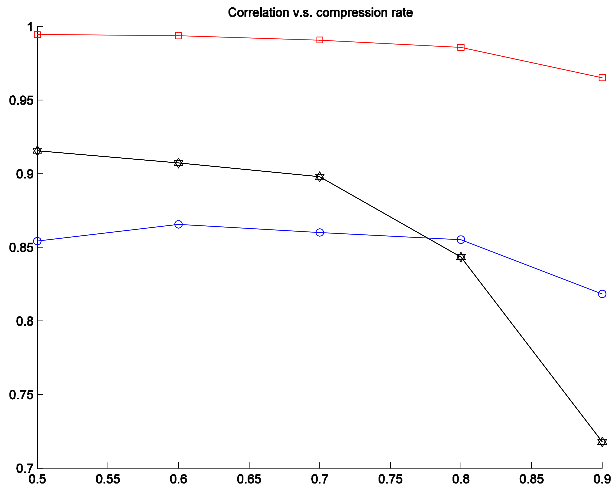



**Fig. 4.** Correlation coefficient versus compression rate of the adaptive sampling algorithm; (circle): curvature vs. inverse target distance; (star): acceleration vs. target based acceleration; (square): tangential velocity vs. target based velocity

# 5 Gesture segmentation, coding and data-driven synthesis

Studies on gesture [10] showed that human gestures can be segmented into distinct phases. Some researches assumed that objective measures can be used to segment hand movement. In particular, Kita et al. showed that abrupt changes of direction, accompanied by a velocity discontinuity indicate phase boundaries in hand trajectories. These observations have been exploited by [11], who proposed a new distance metric to detect phase boundaries, based on the sign of the first and second derivatives of endpoint trajectories. The analysis method described above can be used for automatically segmenting the 3D arm motion. Moreover, it can be used for a compact gesture representation and for data-driven synthesis.

## 5.1 Segmentation

Our segmentation is based on the observation that phase boundaries might occur when the radius of curvature becomes very small, and the velocity decreases at the same time, indicating a change of direction. Our segmentation algorithm is based on the product variable $v(t).\kappa(t)$, and on its approximation, based on the approximated target-based variables : $v_{Tgi}(n_i).\kappa_{Tgi}(n_i)$.

A color-coding method allows to quantify the variations of the variable, according to an equally distribution of its values. The meaning of this coding is presented in table 1.

**Table1.** Coding values for the color coding

| coding | Variable values | Interpretation |
|---|---|---|
| black | --- | lowest values |
| blue | -- | very low values |
| cyan | - | low values |
| green | **0** | average values |
| yellow | + | high values |
| magenta | ++ | very high values |
| red | +++ | highest values |

The color-coding is reported on 3D trajectories, as can be seen in figure 5. When the velocity is very low, the color is green (clear gray). In the contrary, when the velocity is high and the curvature low, the color is red (dark gray). The level of quantification indicates the size of the segmental units. A great similarity can be observed between the segmentation of the curve $v(t).\kappa(t)$ and $v_{Tgi}(n_i).\kappa_{Tgi}(n_i)$ (see figure 5 up and down).

### 5.2 Gesture coding and synthesis

The analysis algorithm described above can be used for representing in a compact way gesture trajectories. These trajectories can be just 3D end-point trajectories, or multidimensional trajectories. In the latter case, the trajectories may be represented by angular postures or Cartesian positions at each joint of the articulated chains.

When applied to 3D end-point trajectories (hand motion), the discrete representation which is provided by the *DPPLA* algorithm can be directly used as input of our motion generation engines. These engines have to determine the angular parameters of the articulated chain, given the end-extremity position and orientation. This operation can be realized through an inverse kinematics (IK) process, such as the GSM controller [12-13] which automatically computes the Euler angles through a non linear optimization approach; we can also use a learning-based inverse kinematics or dynamics scheme [14].

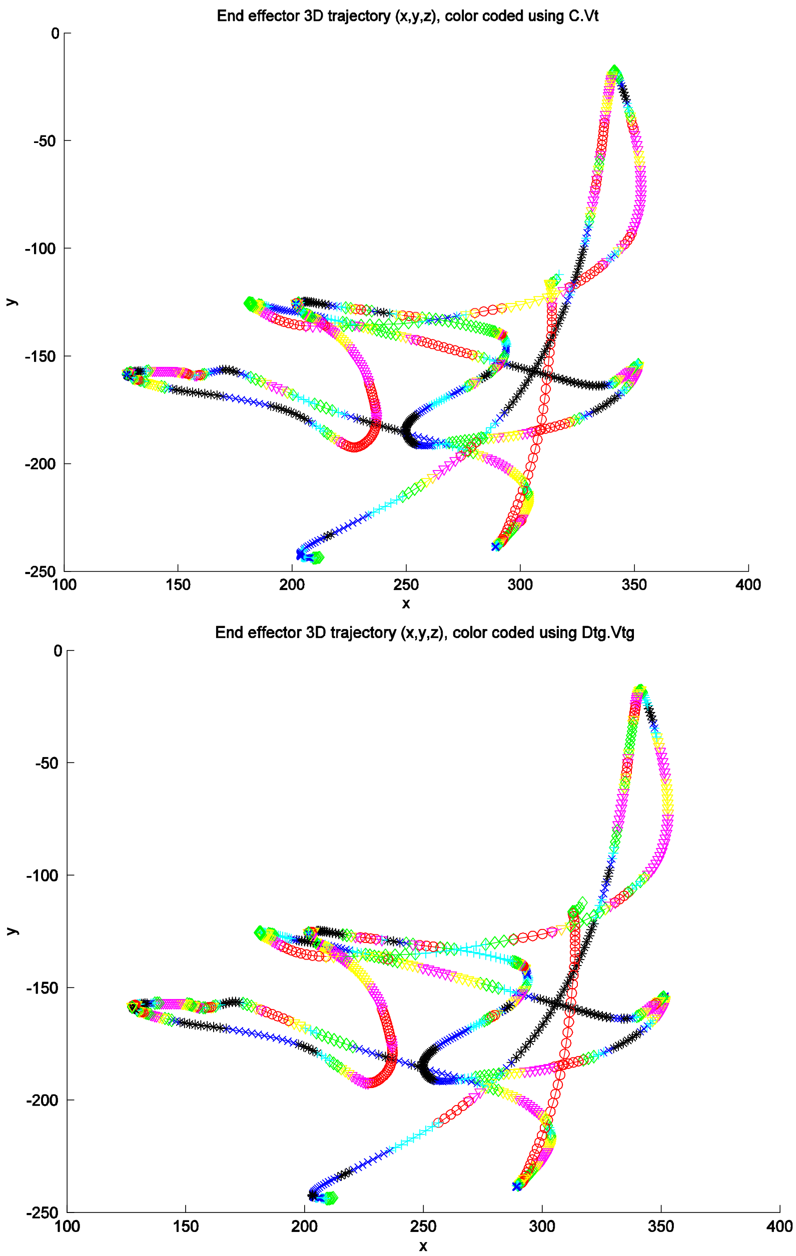



**Fig. 5.** Example of end-point trajectories segmentation (in the *xy* plane) using a color-coding of quantified variables (different gray levels); up: segmentation using the product $\kappa(t).v(t)$; down: segmentation using the product $\kappa_{Tgi}(t).v_{Tgi}(t)$; A great similarity between the two sequences can be observed.

In any case, our inverse methods follow complete end-arm trajectories, or use discrete time-stamped targets localized on the original trajectories. We assume that the target-based representations implicitly contain the main characteristics of the motion style and dynamics. They also provide a way to perform co-articulation, by ensuring transitions between consecutive motion chunks.

The method can be extended to multidimensional trajectories. In the same way, we can represent gesture sequences as target-based vectors evolving with time. The method has been performed for 6D trajectories, representing wrist and elbow Cartesian trajectories. The results are similar to the ones obtained with 3D trajectories: we are able to identify segments along the sequence, with a varying compression factor.

## 6 Conclusion and future work

This paper presented a method for computing an approximation of the curvature and velocity characterizing arm trajectories. This method is applied on compressed data, obtained from an adaptive sampling algorithm. This algorithm extracts discrete target patterns from raw data, for a given compression rate. Given a desired trajectory, we already showed that the targets patterns can represent in an optimal way the original trajectory.

We showed that the target-based approximations are correlated with the instantaneous tangential velocity and curvature. They can therefore be used as an alternative to represent both the shape and the kinematics of end-point trajectories. Moreover, this representation can be adjusted by adapting the compression rate, according to its influence on the correlation. The results obtained for 6D trajectories are very promising. This method for analyzing the shape and kinematics of gesture trajectories may lead to a new analysis tool for multidimensional data.

These empirical approximations provide a significant way to segment gestures. The measure proposed in this paper, in terms of the product of the target-based velocity by the target-based curvature, gives us indeed an original means of delimitating segments which are more or less short, depending on our algorithm parameterization. In order to affirm that these segments represent meaningful components, we should compare them with those obtained through manual segmentation. Anyway, for gestures composed of chunks whose kinematics strongly discriminate them (acceleration, deceleration …), it might by interesting to use our optimized automatic algorithm as an alternative method to the geometrical ones. In future works, an optimal compression rate might be found empirically, by temporally aligning the proposed segmentation with a semantically interpretable segmentation. Other variables should also be tested for segmentation, and confronted to manual segmentation.

Finally, the analysis method defines a possible representation of motion trajectories, based on a kinematic interpretation of the sequences. This representation might be useful for motion retrieval in motion database, or motion synthesis driven by captured data.

## References


1. Marteau, P.F., Gibet, S. Adaptive sampling of motion trajectories for discrete task-based analysis and synthesis of gesture, In Gesture in Human-Computer Interaction and Simulation, GW 2005, Berder Island, France, Revised Selected Papers, Lecture Notes in Computer Science, S. Gibet & al ed., Springer, Vol. 3881, Pages 168-171, 2006.
2. Viviani, P., Terzuolo, C. Trajectory determines movement dynamics. Neuroscience 7:431-437, 1982
3. Soechting, J.F., Terzuolo, C.A. Organization of arm movements in three dimensional space. Wrist motion is piecewise planar. Neuroscience 23:53-61, 1987
4. Chenevière, F., Boukir, S., Vachon, B. A HMM-based dance gesture recognition system. In: Proceedings of the 9th international workshop on systems, signals and image processing, Manchester, UK, June 2002, pp. 322-326
5. Boukir S., Chenevière F.: Compression and recognition of dance gestures using a deformable model, Pattern Analysis and Applications (PAA) Journal, Springer-Verlag, Vol. 7, No 3, (2004) 308-316.
6. Perez J.C., Vidal E.: Optimum polygonal approximation of digitized curves, *Pattern Recognition Letters*, Vol. 15. (1994) 743-750
7. Goodrich M.T.: Efficient piecewise-linear function approximation using the uniform metric. Proceedings of the tenth annual symposium on Computational geometry Stony Brook, New York, United States, (1994) 322 – 331
8. Agarwal P.K., Har-Peled S., Mustafa N.H., Wang Y.: Near-Linear Time Approximation Algorithms for Curve Simplification Proceedings of the 10th Annual European Symposium on Algorithms (2002).
9. Marteau, P.F., Ménier, G., Adaptive multiresolution and dedicated elastic matching in linear time complexity for time series data mining, Sixth International Conference on Intelligent Systems Design and Applications (IEEE ISDA2006), Jinan Shandong, China, 16-18 October, 2006.
10. Kita, S., van Gijn, I., van der Hulst, H. Movement phase in signs and co-speech gestures, and their transcriptions by human coders. Gesture and Sign Language in Human-Computer Interaction, GW 1997, Bielefeld, Germany, Lecture Notes in Computer Science, I. Wachsmuth & al ed., Springer, Vol. 1371, Pages 23-35, 1998.
11. A. Majkowska, V. Zordan, and P. Faloutsos. Automatic slicing for hand and body animations. Eurographics/ ACM SIGGRAPH Symposium on Computer Animation (2006), pp. 1-8, M.P. Cani, J. O'Brien (Ed.)
12. Gibet S., Marteau P.F.: A Self-Organized Model for the Control, Planning and Learning of Nonlinear Multi-Dimensional Systems Using a Sensory Feedback, Journal of Applied Intelligence, Vol. 4. (1994) 337-349
13. Lebourque, T., Gibet, S., A complete system for the specification and the generation of sign language gestures. Gesture-Based Communication in Human-Computer Interaction, GW'99, Gif-sur-Yvette (France), In Lecture Notes in Artificial Intelligence, A. Braffort & al ed., Lecture Notes in Computer Science, Springer, Vol.1739, Pages 227-238, 2000.
14. S. Gibet, P.F. Marteau, F. Julliard. Models with Biological Relevance to Control Anthropomorphic Limbs: A Survey. In International Gesture Workshop on Gesture and Sign Languages in Human-Computer Interaction, Revised selected papers, LNAI, Pages 105-119, London, UK, 2002.