Sythesizing and Modeling Human Locomotion Using System Identification

Wael SULEIMAN, André MONIN and Jean-Paul LAUMOND
LAAS - CNRS
7 avenue du Colonel Roche, 31077 Toulouse Cedex 4, France
{suleiman, monin, jpl}@laas.fr

Abstract—Sythesizing human motion signals is difficult because of its multi-dimensional and nonlinear nature. However, the locomotion is a synchronized motion, which means that these signals are related. Using this property, we propose a new method for modeling the human locomotion and identifying this model. The input signals of our model is the trajectory of pelvis and the outputs are corresponding motion signals of whole human body. To identify this system, we considered it as black-box, for which we propose an adapted method of identification using motion capture. We show that the identified model allows to generate various human locomotion in a fast and efficient way.

I. INTRODUCTION

Locomotion is considered as the basic human motion. However, this motion which appears simple and naive to us, is a challenging topic in many sciences domains such as robotics [1], neuroscience [2], biomechanics [3] and computer graphics [4]. Most of the approaches for synthesizing human locomotion consist in processing a set of motion capture data. Their objective is to provide a control model allowing to perform new not recorded motion. Synthesizing human locomotion is an active research area. Wiley et al [5] use a database of motion capture to synthesize a new motion by classical interpolation motion editing techniques. Pettre and Laumond [6] use a technique inspired from the previous method, the main difference is that their space of interpolation is instantaneous velocities space (control space). Another method for automating gait generation was proposed by Sun et al [7]. This method represents the locomotion into sagittal-plane, which becomes the space of interpolation. Kwon and Shin [8] have proposed a method based on motion modeling in the objective of synthesizing on-line locomotion. Unlike these previous approaches, in this paper we propose a new method to resolve the problem of synthesizing and modeling human locomotion. This method is the result of the comprehensive application of the identification theory and its applications to dynamic system. The identification is one of the most important fields in automatic and control domains [9]. Generally speaking, identification concerns with determining a model of an unknown system subject to measured input-output signals. In our case, the input signals are human captured data using a computer vision techniques. These signals are used to train the model. Our objective is to obtain an input-output control model. The inputs are the trajectory of pelvis in $\mathbb{R}^3$ during locomotion and the outputs are motion signals of the whole human body.

As such the general identification problem is very challenging because of its nonlinearity and multi-dimensional nature. The core contribution of this paper is to show that it can be modeled by linear multivariable systems thanks to a decomposing of the human body structure into simple kinematic chains and using exponential-map parameterization. The advantage of this model is that it is easy to manipulate and its computation complexity is very small, therefore the locomotion controller can be used for interactive applications. Moreover, while classical identification techniques apply for single experiment, this work required extending them to multi-experiments cases.

II. DEFINING MODEL STRUCTURES

A. Black-box model

![Fig. 1. A description of the five open kinematic chains of human 3D articulated structure](image)

We consider the human 3D articulated structure presented in Fig.1. It has 23 joints, each joint has three degrees of freedom (DOF). They are represented by Euler angles. This structure is decomposed into five kinematics chains. Each chain is considered as subsystem, which will be modeled separately.

Fig. 2 shows that the subsystem $S_i$, in reality, contains the physical model with its control unit managed by human’s brain. The physical model is a complex nonlinear multivariable one. It can be obtained from biomechanics studies. However, we do not have any information on the control unit. Therefore we propose to consider the physical model and its control unit aggregate. Such a model can be obtained by black-box
Fig. 2. A description of the model of subsystem $S_i$ related to the chain number $i$

approach [10], [11]. The structure of the model is unknown. It takes the following general form

$$
\begin{align*}
x_{t+1} &= f(x_t, u_t) \\
y_t &= h(x_t, u_t) + v_t
\end{align*}
$$

where $f$ and $h$ are unknown nonlinear functions, $x_t \in \mathbb{R}^n$ is the internal state of the system, $u_t \in \mathbb{R}^m$ is the input signals, $y_t \in \mathbb{R}^p$ is the output signals and $v_t$ is the measurement noise. The noise $v_t$ is considered to be independent of the input signal $u_t$. In general, to identify the model in equation (1), one defines a subclass of parametric nonlinear system [11]. In our case, such task is difficult due to the multivariable nature of the problem and the lack of informations about the model. To solve this difficulty, we use the exponential-map parameterization [12]. This mapping ensures proper manipulation of nonlinear joint-angles quantities by linear multivariable model [13]. Therefore the model becomes

$$
\begin{align*}
x_t &= Ax_{t-1} + Bu_{t-1} \\
y_t &= Cx_t + Du_t + v_t
\end{align*}
$$

where $A$, $B$, $C$, $D$ are the constant system matrices. This classical system identification task can be solved by many standard and robust algorithm [9], such as subspace methods (e. g. N4SID in the MATLAB™ system identification toolbox). Unfortunately this function handles only one data set (in our case, only one motion captured trajectory). As in our case we need an algorithm that can handle many data sets in the same time, we will shown how to adapt the classical PO-MOESP method [14] (see Section (III-A)) to deal with multiple data sets.

B. Exponential-map parameterization

The exponential-map maps a vector in $\mathbb{R}^3$ describing the axis and magnitude of a three DOF rotations to the corresponding rotation. Among the various formulations of the Exp-map [15], we have chosen the classical one proposed in [12]. Let $S^3$ be the set of unit-length quaternion and $SO(3)$ is the subgroup of orthogonal matrices with determinant $+1$ (rotation matrices). In this formulation, we use first a map from $\mathbb{R}^3$ to $S^3$, then the standard quaternion map for conversion to $SO(3).

We can formulate an exp-map from $\mathbb{R}^3$ to $S^3$ as follows

$$
e^{[0,0,0]^T} = [0, 0, 0, 1]^T$$

and for $v \neq 0$

$$
e^v = \left[ \sin(\frac{1}{2} \theta) \hat{v}, \cos(\frac{1}{2} \theta) \right]^T = [q_v, q_w]^T$$

where $q_v$ is the vector part of quaternion, $\theta = |v|$ and $\hat{v} = \frac{v}{|v|}$. This transformation maps $v$ to a unit quaternion representing a rotation of $\theta$ about $\hat{v}$. The inverse of exp-map, the “log” map from $S^3$ to $\mathbb{R}^3$ is given as follows

$$
v = \log(e^v) = \frac{2 \cos^{-1}(q_w)}{|q_v|} q_v
$$

As most of the motion captured data are represented by Euler angles, a conversion from and to Euler angles is necessary. This conversion can be obtained through the matrix of rotation [16].

C. Input/output choice

It is well known that to achieve a good identification, the inputs signals should be persistently excited. Therefore considering the cartesian positions of pelvis during the locomotion directly as input signal hands out a poor model. However, by analyzing the motion signals, we have observed that the three rotations of the pelvis during the locomotion can be assumed as candidates input signals for the chains 1, 2 and 3, and the three rotations of lower neck can be assumed as candidates input signals for the chains 4 and 5 (Fig. 1). This choice is based on the kinematic structure and the frequency spectrum of these signals. Therefore the schema of identification becomes as illustrated in Fig. 3, where $PS$ denotes preliminary subsystem. This subsystem allows to generate the three rotations of pelvis as outputs, then they are transformed into exp-map representation to play the role of input signals to the subsystems $S1$, $S2$ and $S3$. The input signals of $S4$ and $S5$ are the three rotations of lower neck in exp-map representation. These signals can be obtained from the outputs of identified model of $S3$. To validate the above assumptions, we should be able to identify the models, which approximate the outputs of each subsystem (chain) accurately enough.

III. IDENTIFYING MODELS

A. overview of PO-MOESP method

The PO-MOESP method is a class of subspace model identification [14]. Given the linear state space model

$$
\begin{align*}
x_t &= Ax_{t-1} + Bu_{t-1} \\
y_t &= Cx_t + Du_t + v_t
\end{align*}
$$

where $x_t$, $u_t$, $y_t$ and $v_t$ are defined as in Section (II-A). The intent of the method is to calculate an estimation of the quadruple $[A, B, C, D]$. First, the input and output are stocked in Hankel matrices form

$$
U_{1,\alpha,N} = \begin{pmatrix}
  u_1 & u_2 & \cdots & u_{N-\alpha+1} \\
  u_2 & u_3 & \cdots & u_{N-\alpha+2} \\
  \vdots & \vdots & \ddots & \vdots \\
  u_\alpha & u_{\alpha+1} & \cdots & u_N
\end{pmatrix}
$$

1973

Authorized licensed use limited to: LAAS. Downloaded on October 27, 2008 at 13:46 from IEEE Xplore. Restrictions apply.
where the first subscript refers to the index of the first data sample, \( \alpha \) stands for the number of rows in the matrix, and \( N \) refers to the index of the last data sample in the matrix. The number of rows should be chosen to be greater than the system order \( n \) [17]. The output Hankel matrix \( Y_{1,0,N} \) and the noise Hankel matrix \( V_{1,0,N} \) are defined analogously. The following input-output equation is then easily derived from the system description (5):

\[
Y_{1,0,N} = \Gamma_\alpha X_{1,N-\alpha+1} + \Phi_\alpha U_{1,0,N} + V_{1,0,N} \tag{6}
\]

Where \( \Gamma_\alpha \) is the extended observability matrix of the system, \( \Phi_\alpha \) is a block-triangular matrix

\[
\Gamma_\alpha = \begin{bmatrix}
C \\
CA \\
... \\
CA^{(\alpha-1)}
\end{bmatrix}
\]

\[
\Phi_\alpha = \begin{bmatrix}
D & 0 & 0 & ... & 0 \\
CB & D & 0 & ... & 0 \\
CA & CB & D & ... & 0 \\
... & ... & ... & ... & ... \\
CA^{(\alpha-2)}B & ... & CB & D
\end{bmatrix}
\]

and \( X_{1,N-\alpha+1} = [x_1 \ x_2 \ ... \ x_{N-\alpha+1}] \).

PO-MOESP method uses the past input and output data as an instrumental variable to remove the effects of the noise term. Consider the following QR factorization:

\[
\begin{bmatrix}
U_{1+\alpha,0,N} \\
U_{1,0,N-\alpha} \\
Y_{1,0,N-\alpha} \\
Y_{1+\alpha,0,N}
\end{bmatrix} = \begin{bmatrix}
R_{11} & 0 & 0 & 0 \\
R_{31} & R_{22} & 0 & 0 \\
R_{31} & R_{32} & R_{33} & 0 \\
R_{41} & R_{42} & R_{43} & R_{44}
\end{bmatrix} Q^T \tag{7}
\]

where the \( R_{ii} \) are lower-triangular matrices and the column-unitary matrix \( Q (Q^T Q = I) \) is partitioned according to the dimension of lower-triangular matrices \( R_{ii} \) as

\[
Q = [Q_1 \ Q_2 \ Q_3 \ Q_4]
\]

If we consider the quantity

\[
[R_{42} \ R_{43}] = Y_{1+\alpha,0,N} [Q_2 \ Q_3] = (\Gamma_\alpha X_{1,N-\alpha+1} + V_{1+\alpha,0,N}) [Q_2 \ Q_3]
\]

it is not difficult to see that

\[
\lim_{N \to \infty} \frac{1}{N} V_{1,0,N} [Q_2 \ Q_3] = 0 \ w.p.1
\]

where \( w.p.1 \) denotes with probability of one. Consequently, the effects of the noise will vanish, thus \( \Gamma_\alpha \) is extracted from the matrix \( [R_{42} \ R_{43}] \) using a singular value decomposition (SVD) as follows

\[
[R_{42} \ R_{43}] = U_n S_n V_n^T + U_n^\perp S_2 V_n^\perp T \tag{8}
\]

where the matrix \( S_n \) contains the principals singular values (further than threshold), the dimension of this matrix yields \( n \) (the dimension of \( x_1 \)). Then the matrix \( \Gamma_\alpha \) can be estimated as follow

\[
\Gamma_\alpha = U_n S_n^{1/2} \tag{9}
\]

and so we estimate the matrices \( C \) et \( A \) directly from \( \Gamma_\alpha \)

\[
C = \Gamma_\alpha (1 : p_i, :) \\
A = \Gamma_\alpha (p + 1 : p\alpha, :)
\tag{10}
\]

In order to find \( B \) and \( D \), we consider the least-squares solution to the overdetermined system of equations

\[
(U_n^\perp)^T [R_{31} \ R_{32} \ R_{41}] = (U_n^\perp)^T \Phi_\alpha [R_{21} \ R_{22} \ R_{31}] \tag{11}
\]

which provides a consistent estimate of \( \Phi_\alpha \), from which \( B \) and \( D \) are easily calculated.

B. Dealing with Multiple Data Sets

It is clear that the model representation (6) holds for an arbitrary non-zero initial conditions. For that, non-zero initial conditions have no effect at all on the calculations of the quadruple \( [A, B, C, D] \). As a result, dealing with multiple

Fig. 3. An overview of the identification schema
data sets does not introduce an additional problem. As we have mentioned, identifying the models of locomotion involves concatenating multiple data sets corresponding to various trajectory. Therefore we adapt PO-MOESP to deal with multiple data sets as follows. Consider the following data sets
\[ \{u_i, y_{1,i}, y_{2,i}, \ldots, y_{l,i}\}_{i=1}^{N_i} \text{ and } i = 1, 2, \ldots, K \] (12)
where each data set corresponds to a trajectory (experiment). For each input/output data set, we obtain the following data equation
\[ Y_{1,\alpha,N_i} = \Gamma_{\alpha}X_{1,N_i-\alpha+1} + \Phi_{\alpha}u_{t,\alpha,N_i} + v_{t,\alpha,N_i} \] (13)
Then the data equations can easily be combined as follows
\[ [Y_{1,\alpha,N_1} \cdots Y_{1,\alpha,N_k}] = \Gamma_{\alpha} [X_{1,N_1-\alpha+1} \cdots X_{1,N_k-\alpha+1}] + \Phi_{\alpha} [U_{1,\alpha,N_1} \cdots U_{1,\alpha,N_k}] + [V_{1,\alpha,N_1} \cdots V_{1,\alpha,N_k}] \] (14)
The structure of this equation is similar to that of the original data equation (6), for that we still can use the main body of PO-MOESP algorithm by computing QR factorization of the following matrix
\[ [U_{1,\alpha,N_1} U_{1,\alpha,N_2} \cdots U_{1,\alpha,N_k}] [U_{1,\alpha,N_1} U_{1,\alpha,N_2} \cdots U_{1,\alpha,N_k}] \cdots [U_{1,\alpha,N_1} U_{1,\alpha,N_2} \cdots U_{1,\alpha,N_k}] \] (15)
instead of that in equation (7).

C. Estimating initial conditions

The initial conditions are the initial value of internal state \( (x_0) \). Estimating \( x_0 \) is necessary to have a correct initial posture of virtual character. However, this estimation should be done for one data set (one trajectory). Considering a data set \( \{u_i, y_{1,i}\} \), we estimate \( x_0 \) from the following equation
\[ \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{l-\alpha} \end{bmatrix} = \Gamma_{\alpha}x_0 + \Phi_{\alpha} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{l-\alpha} \end{bmatrix} \] (16)
This equation provides an estimation of \( x_0 \) using the pseudo-inverse of matrix.

D. Dealing with stable models

On account of concatenating various data sets, we may be found unstable models. To overcome this problem we should choose a stable model verifying \( \rho(A) < 1 \), where \( \rho(A) \) stands for the spectral radius of \( A \). Lacy et al [18] have proposed a constrained optimization method to ensure the asymptotical stability of the identified model in the context of subspace identification methods. The condition \( \rho(A) < 1 \) is replaced by equivalent Lyapunov inequalities
\[ \rho(A) < 1 \iff \exists P > \delta I_n : P - APA^T \geq \delta I_n \]
where \( \delta > 0 \).

Once we obtain an estimate of the extended observability matrix \( \Gamma_{\alpha} \), the optimization problem is to minimize
\[ J_{\alpha}(A) \triangleq \| (\Gamma_{1,\alpha} - \Gamma_{0,\alpha-1} A) R_{\Gamma} \|^2_2 \] (17)
subject to
\[ P - \delta I_n \geq X^T P \geq \delta I_n \] (18)
where \( \Gamma_{i,j} \in \mathbb{R}^{(i-j+1) \times n} \) is given by
\[ \Gamma_{i,j} \triangleq \begin{bmatrix} C(A)^i \\ C(A)^{i+1} \\ \vdots \\ C(A)^j \end{bmatrix} \]
and \( X \triangleq AP \). We then compute \( \hat{A} \) from \( \hat{X} \) and \( \hat{P} \) as \( \hat{A} = \hat{X} \hat{P}^{-1} \). If we let \( R_{\Gamma} = P \) the problem is converted to an optimization problem that involves minimizing a linear function over symmetric cones (see [18] for more details).

IV. IDENTIFICATION ALGORITHM

To identify the model of locomotion, we should identify the models of subsystems \( PS \) and \( S_i : i = 1, \ldots, 5 \). The schema of identification process is illustrated in Fig. 3. Note that we should consider all data sets in the identification task, therefore we use the extension of PO-MOESP method to deal with multiple data sets. We summarize the identification algorithm as follows

1) Identifying the model of \( PS \): The inputs of this model are the cartesian positions of pelvis \( X = [x_t, y_t, z_t]^T \) and the outputs are the three rotations of Pelvis \( \Theta = \begin{bmatrix} \theta_t^x, \theta_t^y, \theta_t^z \end{bmatrix}^T \). Estimating \( \Theta \) can be done directly from the trajectory of pelvis (i.e. \( \theta_t \) is the tangent angle of the trajectory in the plane \( \{y,z\} \)).

2) Identifying the models of \( S_i \): To solve this problem, we use PO-MOESP method with the corresponding inputs-outputs signals as illustrated in Fig. 3. For example, to identify the subsystem \( S_i \), First, we should define the input-output signals. By using the previous identified model of \( PS \), we can generate the three rotations of pelvis as output, then these rotations are transformed to exp-map representation to play the role of input signals to \( S_i \). The output signals are the exp-map representation corresponding to Euler angles of chain 1 (Fig. 1). Second, we apply PO-MOESP method to estimate the system matrices \( (A, B, C, D) \). Note that as the subsystems are obviously stable, we enforce the identified models to be stable by applying the method explained in Section (III-D).

The above explained algorithm of identification provides a model, which takes as input a real trajectory (a trajectory obtained from motion capture) of pelvis and gives the motion signals of whole human body as outputs. As the implementation of our model will be done on an artificial trajectory \( X_t = [x_t, y_t, z_t]^T \) (e.g. a composition of Bézier curves).
a preprocessing of the trajectory is needed. In fact, a real trajectory can be decomposed into two parts as follows

\[ X_t = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} + \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \hat{X}_t + \hat{X}_t \]  

(19)

where \( \hat{X}_t \) is the main trajectory of pelvis and \( \hat{X}_t \) is related to character zig-zags relative to his trajectory [19]. For that, transforming an artificial trajectory into a real one can be done by identifying \( \hat{X}_t \). Extracting \( \hat{X}_t \) from the real trajectories can be done by filtering \( x_t, y_t \) and \( z_t \) through a low pass filter, so we obtain \( \hat{X}_t \).

As \( \hat{X}_t \) is related to character zig-zags relative to his trajectory, \( \hat{X}_t \) is oscillated signals. Such signals can be modeled by a subspace representation without input sequence. This subclass of subspace representation has the following form

\[ z_t = A z_{t-1} \]

\[ \hat{X}_t = C z_t + v_t \]  

where \( z_t \) is the internal state. To estimate the matrices \( A \) and \( C \), we can use PO-MOESP method. As we know that the model is not stable, we do not enforce it to be stable.

V. EXPERIMENTAL RESULTS

In this section, we represent the results obtained by our method. We have used a database of motion capture (10 different trajectories), which corresponds to one individual character. Fig. 4 represents a description of the implementation of locomotion controller using the identified models. We note that the dimensions of models corresponding to preprocessing, \( PS \) and \( (S_i : i = 1, 2, \ldots, 5) \) subsystems are 10, 10, 15, 15, 10, 10, 10 respectively, and the computing time for modeling and synthesizing was 1250 sec using a PC with 1.5 GHz processor and 512 MByte of RAM. We have validated the obtained model of locomotion on two examples

1) First example: In this example, the trajectory of pelvis has been generated artificially (straight line and arc of circle). Fig. 5 illustrates screenshots of the obtained result.

2) Second example: In order to verify the generality of the identified model, we consider a real trajectory corresponding to another character. This trajectory has been normalized such as the height of pelvis of this character becomes the same of identified character. Note that, in this example the preprocessing subsystem is not used. Fig. 6 illustrates screenshots of the obtained result.

The most observed visible artifact is footskate account for the kinematic constraints imposed by the environment are not included in the identified model. However, such an artifact can be corrected by existing methods if the footplants are annotated [13], [20] as postprocessing. Indeed, this task may violate the real-time constraints, therefore we can neglect this artifact in the case of interactive applications. Videos related to this work are available at http://www.laas.fr/~suleiman.

VI. CONCLUSION

In this paper, we have proposed a new method to model human locomotion. The basic idea is decomposing the articulated structure of 3D human body into five open kinematics chains, then identifying the multivariable model of subsystem corresponding to each chain. The first obtained results are encouraging, therefore enhancing the results to be more robust and realism is the intent of our future work. Finally, we believe that combining the techniques of identification and proven techniques of motion editing has a bright future.

VII. ACKNOWLEDGMENT

The data used in this paper was obtained from mocap.cs.cmu.edu. The database was created with funding from NSF EIA-0196217.

REFERENCES

Fig. 4. An overview of the implementation of the method using the identified models

Fig. 5. First example: screenshots of the application of our locomotion model on an artificial trajectory of pelvis (straight line and arc of circle)

Fig. 6. Second example: screenshots of the application of our locomotion model on a real trajectory of pelvis