REVISTA COLOMBIANA DE COMPUTACIÓN
Volumen 10, número 2
Págs. 60 - 73

Finding the Underlying Structure or Similarity to Others Complex Systems

Hernando Castañeda Marin * Wladimir Rodriguez Graterol † Carlos Arturo Parra Ortega *

Fecha de Recibido: 04/12/2007 Fecha de Aprobación: 02/07/2008

Resumen

La motivación principal en este documento es desarrollar algunos métodos o técnicas que nos permita estudiar sistemas complejos (en el sentido de encontrar su estructura fundamental o su similaridad con otros). Si nosotros poseemos estas técnicas, seremos capaces de abordar una serie de problemas de la vida real que hasta ahora no tienen una solución satisfacible. Ejemplos de tales problemas son el reconocimiento de objetos en 3D, reconocimiento de palabras manuscritas, interpretación de las señales biomédicas y reconocimiento de voz. También presentamos una técnica para analizar los sistemas dinámicos basada en su comportamiento, donde este puede ser determinado a partir de las trayectorias de salida, se utiliza reconocimiento de patrones dinámicos para el análisis de los sistemas. Lo anterior nos permite buscar estructura de datos y su clasificación en categorías de tal forma que la similaridad entre estructuras de la misma categoría sea alta y las de diferente categoría con valores de similaridad baja.

Palabras clave*: Reconocimiento de Patrones Dinamicos, Sistemas Dinamicos , Inteligencia Artificial.*

Abstract

The main motivation of this paper is to develop some methods or techniques that will allow us to study complex systems (in the sense of finding their underlying structure or their similarity to others). If we have these techniques, we will be able to tackle a series of real life problem that until now has had no reliable solution. Examples of such problems are 3D-object recognition, handwritten word recognition, interpretation of bio-medical signal and speech recognition. In this paper, we will present one techniques to analyze dynamical system based on their behavior, where that behavior can be determined from the system output trajectories. We will use dynamic pattern recognition concepts for dynamic system analysis. It allows us to search for structures in data and classify these structures into categories, such that similarity between structures of the same category is high and the similarity between structures of different categories is low.

Keywords: *Dynamic Pattern Recognition , Dynamics System , Artificial intelligent*

* Universidad de Pamplona , Facultad de Ingenieria - CICOM ,hcastaneda@unipamplona.edu.co
† Universidad de los Andes(Ve) ,Faculatd de Ingenieria , Departamento de Computacion , wladimir@ula.ve
* Universidad de Pamplona , Facultad de Ingenieria - CICOM ,cparra@unipamplona.edu.co

1 Introduction

The similarity plays a fundamental role in the theories of knowledge and behavior and has been extensively studied in the psychology literature. Traditionally, dynamic system has being studied using formal mathematical theories.

However, these approaches to system modeling perform poorly for complex, nonlinear, chaotic, and uncertain system. We believe that a possible way to study and analyze such dynamical system is to restate the problem as a similarity problem. We ask "is it possible to find a match or similarity between the dynamic system under study and know dynamical system" this approach is motivated by "Case-Based-Reasoning" where the premise is that once a problem has been solved, it is often more efficient to solve a similar problem by starting from the old solution, rather than rerunning all the reasoning that was necessary the first time.

2 Problem Formulation

Traditional approaches to analysis system –e.g. trying to find a mathematical model that describe the output as a function of state variable and the input perform poorly when dealing with complex system. This may be due to their nonlinear, time-varying nature to uncertainty in the available measurement.

We can approach the analysis of dynamic system in two different ways: the first is based on the existence of a state measuring mechanism in the form of mathematical model, in the absence of such a measuring mechanism we must resort to some perceptual mechanism, that allow us to perceive the underling structure of the system, based on the behavior of the dynamic system. The similarity measure is one the possible perceptual mechanism that can be used to analyze such systems. One off motivations of this dissertation is to discover ways to use structural similarity as mechanics to study dynamic system.

2.1 Description Mathematical and Modeling Dynamic System

The classic methods of recognition of patterns should be tuned to consider desirable problems from the dynamic point of view, that is to say the process of objects are described with sequences of temporary observations.
In the design of dynamic systems and the analysis in the domain of the time the concept of states of a system is used, a dynamic system is

usually modeled by a system of differential equations. To obtain dynamic system by differential equations that represent the relationship between the input variables $u_1(t)$, $u_2(t),...u_p(t)$ and the output variables $y_0(t)$, $y_1(t),...y_q(t)$, the intermediate variables receive the name of state variables $x_1(t)$, $y_2(t),...x_n(t)$. A group of state variables in any instant determines the state of the system in this time.

If the current state of a system and the value of the variables are given for $t > t_0$, the behavior of the system can be described clearly.

The state of the systems is a set of real numbers in such a way that the knowledge of these numbers and the values of the input variables provide the future state of the system and the values of the output variables by the equations that describe the dynamics of the system. The state variables determine the future behavior of the system when the current state of the system and the values of the entrance variables are known [2].

The multidimensional space of observation induced by the state variables receives the name of space of states. The solution of a system of differential equations can be represented by a vector *x(t)* that corresponds to a point in the state space in an instant of time *t*. This point moves in the space of states like steps of time. The appearance or on the way to this point in the space of states is known like a trajectory of the system. For an initial state and end state given (1) a number infinite of input vectors exist that correspond to trajectories with start and end points. Of another side considering any point in the state space goes a trajectory exactly by this point [3].

Considering dynamic systems in the control theory, a lot of attention has been paid to the adaptive control. The main reason to introduce this investigation area is the one of obtaining controllers that its parameters can adapt to the changes in the process dynamic and perturbation characteristic.

2.1.1 Linear Dynamic Systemss Class

The state of transition of the dynamic system in the internal space and the mapped from the space of internal states to the space of observations is modeled by means of the following lineal equations.

$$\begin{aligned} x_t &= F^{(i)} x_{t-1} + g^{(i)} + w_t^{(i)} \\ y_t &= H x_{t-1} + v_t \end{aligned} \qquad (1)$$

Where $F^{(i)}$ is a transition matrix; $G^{(i)}$ is a bias vector. *H* Is a transition matrix that defines the lineal projection from a space of internal state to the observation space, Notice that each dynamic system has, $F^{(i)}$ $G^{(i)}$ y $W_t^{(i)}$ individually. It is assumed that each $W^{(i)}$ is noise identifier and v has normal distribution. $N_{x1}(0,Q^{(i)})$ and $N_y(0,R)$ respectively.

The classes of dynamic systems can be categorized by the eigenvalues of the transition matrix which determines answers of input zero of the system. With other words, these eigenvalues determines the general behavior of patterns (trajectories) with temporary variation in the space of states.

For the concentration of the temporary evolution of the state in the dynamic system, it is assumed that the bias term and that process of noise is zero in the equation (1). Using the decomposition of the eigenvalues in the transition matrix the following expressions are obtained:

$$F = E\Lambda E^{-1} = [e_1,...,e_n]\, diag\,(\lambda_1,...,\lambda_n)[e_1,...,e_n]^{-1} \quad (2)$$

It can be solved the state in the time x_0 this way:

$$x_t = F^t x_0 = (E\Lambda E^{-1})^t x_0 =$$

$$E\Lambda^T E^{-1} x_0 = \sum_{p=1}^{n} \alpha_p\, e_p\, \lambda_p^t \quad (3)$$

$$[\lambda_1,\ldots,\lambda_n]^T = E^{-1} x_0$$

Where λ_p *and* e_p they are the corresponding eigenvalues and eigenvectors. The pondered value α_p is determined from the initial state x_0 by the determination of in the complex plane.

Of here, the general patterns of a system can be categorized by means of the position of the eigenvalues (poles) in the complex plane. The determination of the oscillatory states this certain one for their (angle) arguments according to the following rules.

It oscillates if at least an eigenvalue is negative or complex. It doesn't oscillate if all the eigenvalues has real numbers. The absolute value of the eigenvalue determines the convergence state or divergence in the way: It diverges if at least a value of the eigenvalue exceeds in one. It converges if all the absolute values of the eigenvalue are less than one.

In the table 1 are illustrated states of trajectories with two-dimensional states. The systems can generate patterns for temporary variation if and alone if, this pattern converges to zero. (Using control terms is said that the system this stable one).

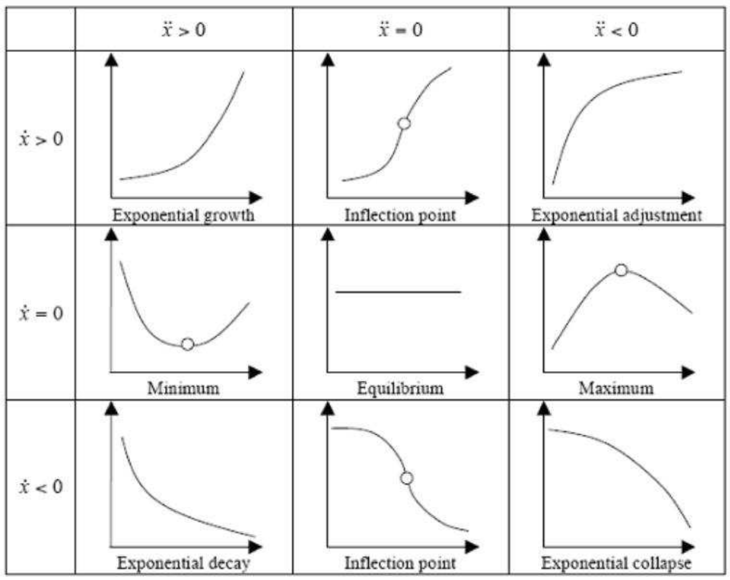


Table 1. Examples of class of dynamic with bidimensionalidad in the phase state

2.1 Dynamic Pattern Recognition Base

Consider a complex system that assuming different states in the course of the time. Each state of the system in the instants of time is considered an object to classify. If a dynamic system is observed temporarily, the variable value of the features constitutes dependent functions of the time.

However each object is not only described by one vector of features in any instant but also for the history of the temporary development of this vector of features.

The objects receive the name of dynamic whether they represent measurements or observations of a dynamic system and it contains history of their temporary development. That is to say each dynamic object is a temporary sequence of observations that is described by a discreet function in the time. The dependent function of the time is represented by a trace, or trajectory, for each object from its initial state to its current state in the space features.

In the figure 1 are illustrated trajectories of dynamic objects in a space of two-dimensional features. If the form of the trajectories is choice as the criterion of similarity the trajectories has three cluster of objects, it can be distinguished {*A, C*}, {*B, D, E, G*} y {*F, H*}. These three clusters is different to those that are recognized as static objects in an instant in the time [6].

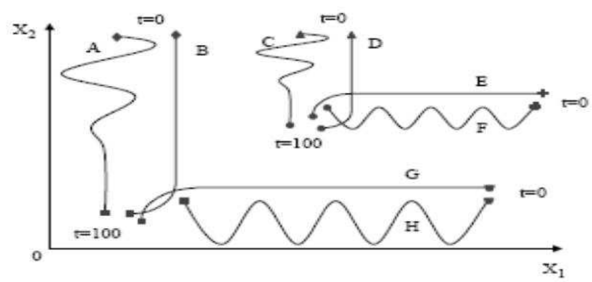


Fig.1. Objects in the two-dimensional space of features

If the form and orientation of the trajectories is chosen then as similarity approach the objects *B, D, E, G* they cannot be considered similar and they are divided in the following two groups {*B, D*} and {*E, G*}. If the form and orientation of the trajectories are considered irrelevant but their closeness space is then a base for a similarity definition, four clusters they are recognized this way {*A, B*}, {*C, D*}, {*E, F*} y {*G, H*}.

This example illustrates the method of classic recognition patterns in the environment dynamic, since it doesn't consider the temporary behavior of the system in consideration [6].

2.2.1 Similarity Measure Based on Characteristic of Trayectories

Consider a complex system that assuming different states in the course of the time. Each state of the system in the instants of time is considered an object to classify. If a dynamic system is observed temporarily, the variable value of the features constitutes dependent functions of the time.

In the previous section we consider a criterion for the comparison between two trajectories, two similarity types among trajectory can considers:

Pointwise Similarity: the smaller Pointwise distance between two trajectories in feature space.

Structural Similarity: the better two trajectories match in form, evolution, characteristic, and the greater the similarity between two trajectories.

For the determination of the similarity structural it is specified relevant aspects of the behavior of the trajectories depending on a concrete application. Based on physical properties of the trajectories (e.g. slope, curvature, smoothness, position and value of inflection points) can be selected, which are then used as comparison criteria.

In such a way, the similarity structural is suited to situation for when we look at the particular patterns in trajectories that should be well matched.

2.2.2 Similarity Structural Based of Slope and Curvature Trajectories.

The curvature of the trajectories of each point describes the grade with the one which a trajectories are s bent in this point. This is evaluated by the coefficient of second derivative of a trajectory in each point that can be defined by the following equation (for a trajectory one-dimensional).

$$cv_k = \ddot{x}_k = \frac{\dot{x}_k - \dot{x}_{k-1}}{t_k - t_{k-1}} \quad k = 3, \ldots, p \tag{4}$$

Where $\dot{x}_k$ denotes the coefficient of the first derived in the point x_k and given for:

$$x_k' = \frac{x_k - x_{k-1}}{t_k - t_{k-1}} \quad \text{k=2,...,p} \tag{5}$$

Substituting the previous equation in the equation of the bend, you arrive to the following equation based on the values of the original trajectories

$$cv_k = \ddot{x}_k = \frac{x_k - 2x_{k-1} + x_{k-2}}{(t_k - t_{k-1})^2} \quad k = 3, \ldots, p \tag{6}$$

The distinctive characteristic when it is considered the curvature, it is the sign of the coefficient of second derivate. If the coefficient is positive in certain period of time, then the trajectory is convex in the interval (near to the end). If the coefficient is negative in certain period of time, the trajectory is concave (near to the low point). If the coefficient is similar to zero in some point that inflection point is called, bend is not presented in this point.

In a trajectory, they can be distinguished seven types of segments (tendencies), each one of those which this characterized by a constant sign in the first one and second derived. Such a triangular representation of tendencies provides qualitative characteristic for a description of segments.

To derive quantitative information starting from the segments, these they are described by the following group of parameters *t(a); t (b)* they are the instants of initial and final time of the segment. See figure 2.

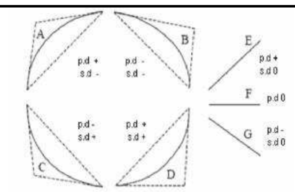
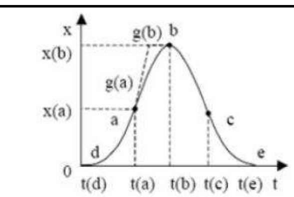
Primitives pattern	Parameters in some segment
	

Fig.2. Qualitative and Quantitative Temporary Characteristics obtained by Segmentation

3 Problem Solution

The key idea for the learning process is that the estimate process is divided in to two stages: the process of clustering of dynamic systems to estimate a group of dynamic systems and a refinement of parameters of the estimated dynamic systems. In this brief, a simple approach is proposed to analyze the dynamic clustering in a trajectory for linear dynamic system.

3.1 Clusters of Trajectories as Dynamic Systemas (Dynamic Objects)

The key idea for the learning process is that the estimate process is divided in to two stages: the process of clustering of dynamic systems to estimate a group of dynamic systems and a refinement of parameters of the estimated dynamic systems.

This is the first stage of the process under consideration; it consists in finding a set of dynamic systems, the number of dynamic systems and their parameters. A group of typical sequences is used (for example a subset of given training data) and the sequences have already been mapping in the space of internal states. The clustering technique estimates a group of dynamic systems; then an estimate is made of the *N* number of dynamic systems and an approximation of the parameters θ_i ($i = 1,..., N$) of the dynamic system.

The second stage is a process of refinement of the parameters of the system based on the algorithm EM. The process is applied to the whole of the given training data, while the clustering process is applied to a select group of training data.

It is assumed that a sequence of many variables y^{t}_{1} □ $y_1,...,y_T$ is a typical training data. The order of the transitions of the dynamic systems won't be considered. You can consider a single set of data of training without losing generality in this stage of the clustering.

A group of dynamic systems simultaneously can be identified $\{D_1, D_2, ..., D_N\}$ for example the number of dynamic systems and their parameters $\theta_1,..., \theta_N$ for some interval groups *I* (For example to segment and to label the sequence) from the sample of training. y^{t}_{1}, Where the number of intervals *k* is also ignored.

It is assumed that the observation matrix *H* is given. Using an of identification technique of dynamic subspaces for example OM95 before beginning the clustering stage, you can simultaneously estimate the observation matrix *H* and their corresponding sequence of internal states x^{t}_{1} □ $x_1,...,x_T$; tit can be mapping in an observation sequence given by the matrix.

3.2 System Identification Bases an Eigenvalues

To identify the parameters with a small set of data training, one has to make restrictions in the eigenvalue to estimate desirable dynamic systems.

This restriction is based on the dynamic stability; the key idea of estimating dynamic stability to give constraint in the eigenvalue. State of dynamics system change in a stable manner if all the eigenvalues are smaller that one

The identification of the system without restrictions is conditioned so that the temporary range [*b, e*] is represented by the lineal dynamic system D_i.

The transition matrix $F^{(i)}$ and the vector of bias $g^{(i)}$ of the sequence of internal states $x_b^{(i)},..., x_e^{(i)}$ are considered. This problem of estimate of parameters becomes a problem of minimization of prediction of errors.

This vector of error prediction can be determined by starting from equation (1) and estimating the matrix $F^{(i)}$ and the vector of biases $g^{(i)}$. Their formulation is:

$$\varepsilon_t = x_t^{(i)} - \left(F^{(i)} x_{t-1}^{(i)} + g^{(i)}\right) \quad (7)$$

So the sum of the norms of the squares of all the error vectors in the range [*b, e*] becomes:

$$\sum_{t=b+1}^{e} \left\|e_t\right\|^2 = \sum_{t=b+1}^{e} \left\| x_t^2 - \left(F^{(i)} x_{t-1}^{(i)} + g^{(i)}\right) \right\|^2 \quad (8)$$

Finally the optimal values of $F^{(i)}$ and $g^{(i)}$ by the solution can be determined by solving the following problem of the least mean square.

$$F^{(*i)}, g^{(*i)} = \arg\ \min \min_{F^{(i)}, g^{(i)}} \sum_{t=b+1}^{e} \left\|e_t\right\|^2 \quad (9)$$

The identification system with restrictions in the eigenvalue of the transition matrix $F^{(i)}$ is deduced from the estimated transition matrix and the estimated vector bias and has the following form:

$$\begin{aligned} F^{(i)*} &= \widehat{X}_1^{(i)} \widehat{X}_0^{(i)+} \\ g^{(i)*} &= m_1 - F^{(i)*} m_0 \end{aligned} \quad (10)$$

Where m_0 and m_1 are the vectorial means of the columns in $x_0^{(i)}$ respectively**.** The temporal interval [*b, e*] is represented by a linear dynamic system D_i. Thus we can estimate the transition matrix $F^{(i)}$ by the following equation:

$$\begin{aligned} F^{(i)*} &= \operatorname{argmin}_{F^{(i)}} \left\| F^{(i)} X_0^{(i)} - X_1^{(i)} \right\|^2 \\ F^{(i)*} &= \lim_{\delta^2 \to 0} \left[X_1^{(i)} X_0^{(i)T} \left(X_1^{(i)} X_0^{(i)T} + \delta^2 I \right)^{-1} \right] \end{aligned} \quad (11)$$

Where I is the unit matrix, δ is a positive real value.

In the eigenvalues constraint, the limit is detained before $(X_1^{(i)} X_0^{(i)T}+\delta^2 I)^{-1}$ converges to the pseudo-inverse matrix of $X_0^{(i)}$ [1]. Using Gershgorin`s theorem in linear algebra [4], we can determine the upper bound of the eigenvalue in the matrix.

Suppose $f_{(u,\ v)}^{(i)}$ is an element in row u and column v of the transition matrix $F^{(i)}$. Then, the upper bound of the eigenvalue is determined by $B = \max_u \sum_{v=1}^{n} \left| f_{(u,v)}^{(i)} \right|$. Therefore, we search for a nonzero value for δ; which controls the scale of elements in the matrix $F^{(i)}$, which should satisfy the equation , $B=1$ via iterative numerical methods.

For the evaluation, we used two simulated sequences of a physic problem as training data to verify the clustering method, because it will provide new paper system identification.

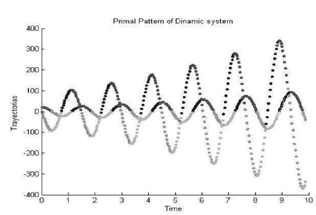
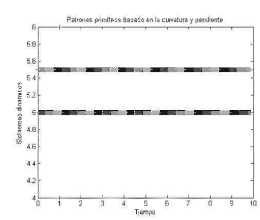
<table>
<tr><th>Dynamic System Behavioral</th><th>Primal Pattern of Dynamic System</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Table 2. Behavior of Sequences Pattern

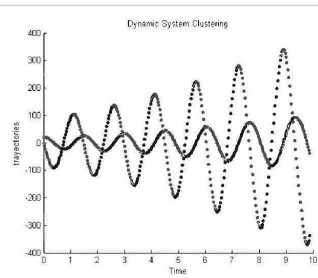
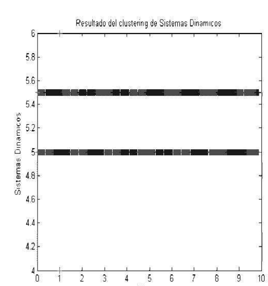
<table>
<tr><th>Dynamic System Clustering</th><th>Two Linear Dynamic System Interval</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Table 3. Clustering Trajectories of Dynamic System

Interval Red and means value	Interval Blue and means value	Deviation of initial points
[97,120]	[74,96]	
m_0 =0.8394 m_1 =–2.49	m_0 =0.7497 m_1 =3.83	-31 34 -33 37
[85,108]	[61,84]	
m_0=0.7591 m_1=–12.41	m_0 =–0.8892 m_1 =10.68	-185 135 -128 133

Table 4 Initial values of Clustering Trajectories of dynamic system

The analysis of the cases: the worst case $\delta^2 = 0$, the best case $\delta^2 = 1$ and the average 0.25 produce following result:

With a data matrix for each dynamic system, the transition matrix is possible to estimate follow way:

$$X^{(1)} = \begin{bmatrix} -31| & -118 \\ -33 & -128 \end{bmatrix} \quad X^{(2)} = \begin{bmatrix} 34 & 135 \\ 37 & 133 \end{bmatrix}$$

The analysis of the cases: the worst case $\delta^2 = 0$, the best case $\delta^2 = 1$ and the average 0.25 produce following result:

$$X^{(1)+} = \left[\begin{bmatrix} -31| & -118 \\ -33 & -128 \end{bmatrix}\begin{bmatrix} -31 & -33 \\ -118 & -128 \end{bmatrix} + \delta^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right]^{-1}$$

$$X^{(2)+} = \left[X^{(2)} = \begin{bmatrix} 34 & 135 \\ 37 & 133 \end{bmatrix}\begin{bmatrix} 34 & 37 \\ 135 & 133 \end{bmatrix} + \delta^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right]^{-1}$$

In worst case:

$$F^{(1)*} = \left[\begin{bmatrix} 14885 & -14081 \\ -14081 & 17473 \end{bmatrix} x10^{-3} \begin{bmatrix} 0.2827 & 0.2278 \\ 0.2278 & 0.2408 \end{bmatrix}\right]$$

$$F^{(1*)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$F^{(2)*} = \left[\begin{bmatrix} 19381 & 19213 \\ 19213 & 19058 \end{bmatrix} \begin{bmatrix} 0.2827 & 0.2278 \\ 0.2278 & 0.2408 \end{bmatrix}\right]$$

$$F^{(2*)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

In average case:

$$F^{(1)*} = \left[\begin{bmatrix} 14885 & 16127 \\ 16127 & 17473 \end{bmatrix} * \begin{bmatrix} 3.1758 & -2.9312 \\ -0.0859 & 2.7054 \end{bmatrix}\right]$$

$$F^{(1*)} = \begin{bmatrix} 0.9388 & 0.0564 \\ -0.0612 & 1.0564 \end{bmatrix}$$

$$F^{(2)*} = \left[\begin{bmatrix} 19381 & 19213 \\ 19213 & 19058 \end{bmatrix} * \begin{bmatrix} 0.0852 & -0.0859 \\ -0.0859 & 0.0866 \end{bmatrix}\right]$$

$$F^{(2*)} = \begin{bmatrix} 1.0002 & -0.0002 \\ 0.0002 & 0.9998 \end{bmatrix}$$

In best cases:

$$F^{(1)*} = \left[\begin{bmatrix} 14885 & 16127 \\ 16127 & 17473 \end{bmatrix} * \begin{bmatrix} 3.1758 & -2.9312 \\ -2.9312 & 2.7054 \end{bmatrix}\right]$$

$$F^{(1)*} = \begin{bmatrix} 0.7588 & 0.2226 \\ -0.2412 & 1.2226 \end{bmatrix}$$

$$F^{(2)*} = \left[\begin{bmatrix} 19381 & 19213 \\ 19213 & 19058 \end{bmatrix} * \begin{bmatrix} 0.0852 & -0.0859 \\ -0.0859 & 0.0866 \end{bmatrix}\right]$$

$$F^{(2)*} = \begin{bmatrix} 1.0007 & -0.0008 \\ 0.0007 & 0.9992 \end{bmatrix}$$

4 Conclusión

In this paper, we proposed a novel computational model, named clustering based on structural similarity to model dynamic systems and their structures. The temporal segmentation and system identification problem need be resolved simultaneously; we showed that the system can analyze dynamic features based on the timing structure extracted from temporal intervals. We applied the proposed model to describe dynamic structure that consists on primitive's pattern. Problem to determine the appropriate number of dynamic systems, there are several well know criteria between find knee of the log -likelihood curve and an evaluation functions that consist in the log-likelihood scores and the numbers free of parameters.

References

[1] Courrieu, P., Fast Computation of Moore-Penrose Inverse Matrices, Neural Information Processing - Letters and Reviews, Vol.8, No.2, August 2005.

[2] Dorf, R.C, Modern Control System, Addison –Wesley, Reading, Mass., 1992.

[3] Föllinger, O., Franke, D., Einführung in die Zustandsbeschreibung dynamischer Systeme. Oldenbourg Verlag, München, 1982.

[4] Franze G., Carotenuto, L., Muraca, P. , On the stability of interval matrices, Proceeding of the 2004 American Control Conference, Boston, Massachusetts June 30 - July 2,2004.

[5] Kawashima, H., Matsuyama, T. , Proc. 3rd International Conference on Advances in Pattern Recognition (S. Singh et al. (Eds.): ICAPR 2005, LNCS 3686, Springer), pp. 229-238, 2005.

[6] Mikenina,Geb, Angstenberger , Larisa, Dynamic Fuzzy Pattern Recognition, Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online ver,2000.