

Identification of Dynamic Complex Trajectories using Gershgorin's Theorem in Principal Component Analysis

Identificación de la Dinámica de Trayectorias Complejas usando el Teorema de Gershgorin en Análisis de Componentes Principales

Hernando Castañeda Marín PhD¹, Wladimir Rodríguez Graterol, PhD², Eliézer Colina Morles PhD³.

1 Universidad de Pamplona, Doctorado en Ciencias Aplicadas ULA, Venezuela

2 Facultad de Ingeniería, Departamento de Computación ULA, Venezuela

3 Facultad de Ingeniería, Departamento de Automatización y Control, Venezuela

hcastaneda@ula.ve; wladimir@ula.ve; ecolina@ula.ve

Recibido para revisión 15 de Marzo de 2008, Aceptado 19 de Mayo de 2008, Versión final 23 de Mayo de 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 satisfactoria. 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 basado en su comportamiento, donde este puede ser determinado a partir de las trayectorias de salida, se utiliza reconocimiento de patrones dinámicos y componentes principales para el análisis de los sistemas. El propósito es encontrar vectores básicos utilizando similaridad estructural tal que la secuencias de las variables de estado puedan ser segmentadas y cada segmento pueda ser descrito como una combinación lineal de vectores básicos.

Palabras clave: Reconocimiento de Patrones, Dinámicos, Sistemas Dinámicos, 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 and principal component analysis concepts for dynamic system analysis. The purpose is to find vectors basis using such similarity structural where the sequences of the state variables can be segmented and each segment can be described as a lineal combination of vectors basis.

Keywords: Dynamic Pattern Recognition, Dynamic System, Artificial Intelligent

I. 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 been 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.

II. 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 underlying 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.

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 $u_1(t), u_2(t), \dots, u_p(t)$ and the output variables $y_0(t), y_1(t), \dots, y_q(t)$, the intermediate variables receive the name of state variables $x_1(t), x_2(t), \dots, x_n(t)$. 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 [3].

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 that corresponds to a point in the state space in an instant of time. 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 [4].

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.

B. 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.

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

Where $F^{(i)}$ is a transition matrix; $g^{(i)}$ is a bias vector. 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 has normal distribution $N_x(0, Q^{(i)})$ and $N_y(0, R)$ respectively.

The classes of dynamic systems can be categorized by the eigenvalue of the transition matrix which determines answers of input zero of the system. With other words, these eigenvalue 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 of process of noise are zero in the equation (1). Using the decomposition of the Eigenvalue in the transition matrix the following expressions are obtained:

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

It can be solved the state in the time with conditions initials this way:

$$\begin{aligned} x_t &= F^t x_0 = (E \Lambda E^{-1})^t x_0 \\ E \Lambda^t E^{-1} x_0 &= \sum_{p=1}^n \alpha_p e_p \lambda_p^t \\ [\lambda_1, \dots, \lambda_n]^T &= E^{-1} x_0 \end{aligned} \tag{3}$$

Where λ_p and e_p they are the corresponding eigenvalue and own vectors. 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 Eigenvalue (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 eigenvalue 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).

III. 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 objects in a space of two-dimensional features. If the form of the trajectories is choice as the criterion of similarity the trajectories then 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]

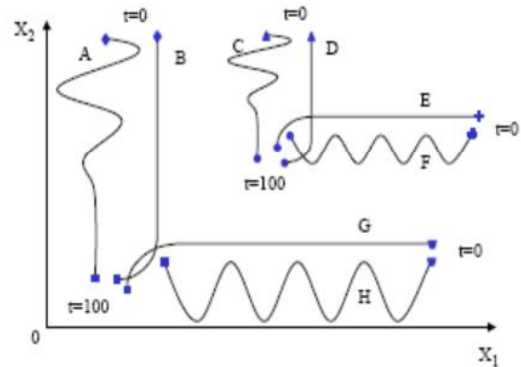


Figura1 Objects in the two-dimensional space of features

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 [7].

A. Similarity structural based of slope and curvature trajectories.

The curvature of the trajectories of each point describes the grade with the one which a trajectory this 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).

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

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

$$\dot{x}_k = \frac{x_k - x_{k-1}}{t_k - t_{k-1}} \quad k=2, \dots, p \tag{5}$$

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

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

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

Table 1 Examples of class of dynamic with bidimensional in the phase state

	$\ddot{x} > 0$	$\ddot{x} = 0$	$\ddot{x} < 0$
$\dot{x} > 0$	 Exponential growth	 Inflection point	 Exponential adjustment
$\dot{x} = 0$	 Minimum	 Equilibrium	 Maximum
$\dot{x} < 0$	 Exponential decay	 Inflection point	 Exponential collapse

to zero in some point that inflection point is called, bend is not presented in this point.

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

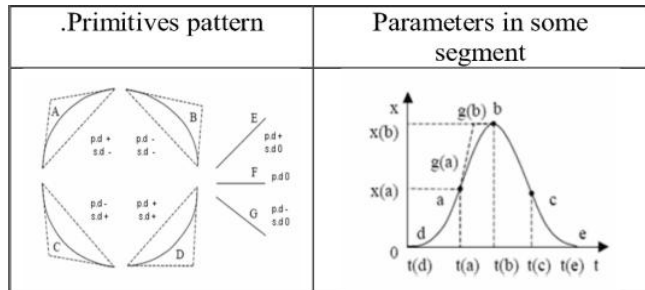


Figure 2 Qualitative and Quantitative Temporary Characteristics obtained by Segmentation

In a trajectories (see figure 3), we can be distinguished four of the 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 characteristic qualitative for a description of the segments.

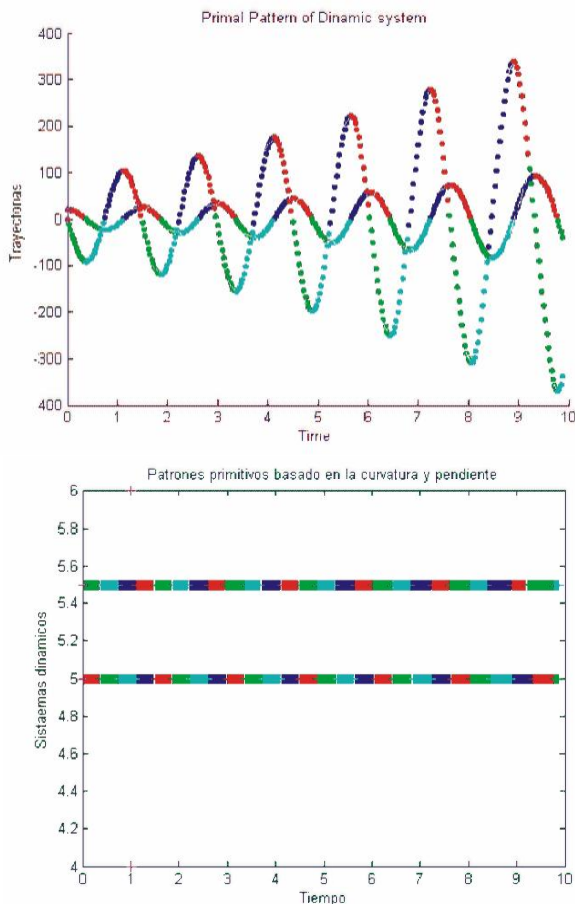


Figure 3, Behavior of Sequences Pattern

IV. PROBLEM SOLUTION

PCA is coupled by eigenvalue decomposition on the correlation matrix or by singular value decomposition on the data matrix. The PCA algorithm computes a Matrix of rank and the decomposition of into the orthogonal basis of size m, such that

$$R = \arg \min_{\substack{X \\ \alpha_i \in (s_i, s_{i+1})}} \|X - RQ\|_F \quad (7)$$

Where $X = x_1(t), x_2(t), \dots, x_n(t)$ he basis vector $v_1(t), v_2(t), \dots, v_m(t)$ are the eigenvectors of the correlation matrix or the right singular vectors of the data matrix respectively.

Given $S = \langle s_1, s_2, s_3, s_4 \rangle$ partitions of the sequence X explicitly as $x_1(t), x_2(t), \dots, x_n(t)$ in continuous intervals where each point $x_i \in X$ belongs exactly to an interval then let us consider a matrix R of the 4 labeling observations (columns) s_i with the colors like they are shown in figure 3 and the size of the rows m corresponding to the biggest size in the sequence of observations, so R is an $n \times d$ matrix. Specifically it has that $R = (r_1, r_2, \dots, r_n)$ where $r_i \in R^d$, for all $i = 1, 2, \dots, n$.

For the estimate of the structure of the dynamic system, we need for each segment $s_i, \alpha_i, \tau \in F$ for $t = 1, \dots, m$, the coefficients of a lineal combination of basic vectors $v_j, j = 1, \dots, 4$ in notation matrix, $A = (a_{ij})$ is a $4 \times m$ matrix of coefficients all observation in segment s_j are represented by the follow single vector:

$$u_j = \sum_{i=1}^m a(j, i) v_i \quad (8)$$

System Identification Bases an Eigenvalue

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 eigenvalue 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 $x_b^{(i)}, \dots, x_e^{(i)}$ 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:

$$\varepsilon_t = x_t^{(i)} - (F^{(i)} x_{t-1}^{(i)} + g^{(i)}) \quad (9)$$

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

$$\sum_{t=b+1}^e \|e_t\|^2 = \sum_{t=b+1}^e \|x_t^{(i)} - (F^{(i)} x_{t-1}^{(i)} + g^{(i)})\|^2 \quad (10)$$

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.

$$F^{(i)*}, g^{(i)*} = \arg \min_{F^{(i)}, g^{(i)}} \sum_{t=b+1}^e \|e_t\|^2 \quad (11)$$

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:

$$\begin{aligned} F^{(i)*} &= \hat{X}_1^{(i)} \hat{X}_0^{(i)+} \\ g^{(i)*} &= m_1 - F^{(i)*} m_0 \end{aligned} \quad (12)$$

Where m_0 and m_1 are the vectorial means of the columns in $X_0^{(i)}$ and $X_1^{(i)}$ 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:

$$\begin{aligned} F^{(i)*} &= \arg \min_{F^{(i)}} \|F^{(i)} X_0^{(i)} - X_1^{(i)}\|^2 \\ F^{(i)*} &= \lim_{\delta \rightarrow 0} [X_1^{(i)} X_0^{(i)T} (X_1^{(i)} X_0^{(i)T} + \delta^2 I)^{-1}] \end{aligned} \quad (13)$$

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

In the eigenvalue 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

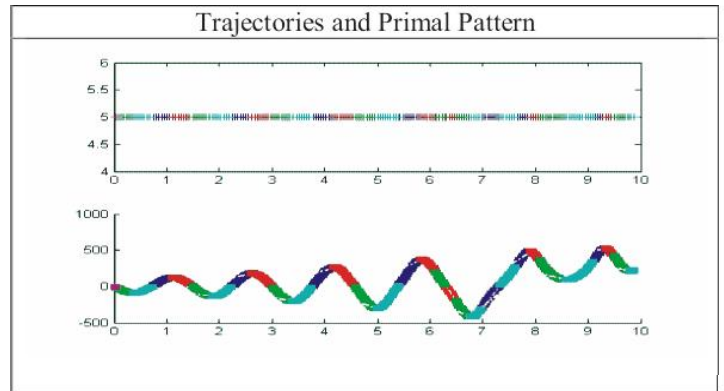
eigenvalue is determined by $B = \max_u \sum_{v=1}^n |f_{(u,v)}^{(i)}|$

. 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.

V. RESULTADOS

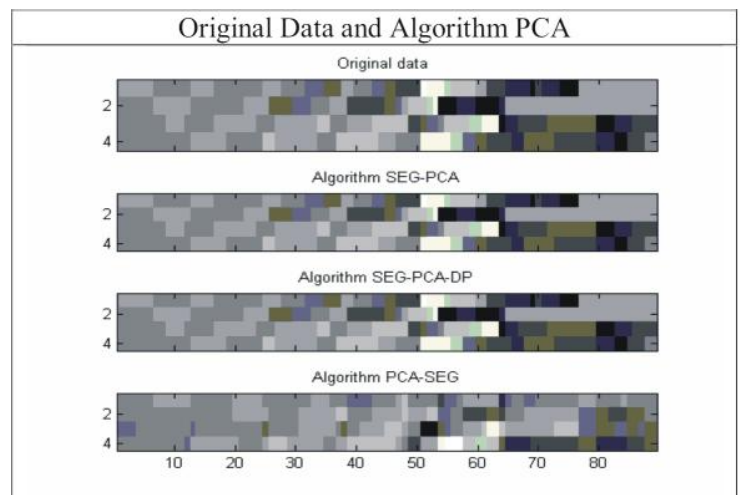
For the evaluation, we used one simulated sequences of a physic problem (pendulum simple) as training data to verify the clustering method, because it will provide new paper system identification:



We use three algorithms, they are as follow (see table 2):

- Seg-PCA: first segment into segments, then find a basis of size for the segment means.
- Seg-PCA-DP: first segment into segments, then find a basis of size for the segment means, and then refine the segmentation by using the discovered basis.
- PCA-Seg: First do PCA to dimension then obtain the optimal segmentation of the resulting -dimensional sequence.

Table 2 Original Data and Algorithm PCA



The matrix eigenvalue problem is one of the central problems in the numerical linear algebra. It is formulated as follow:

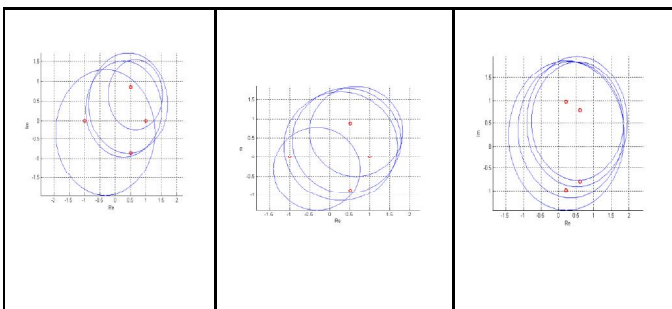
Given a square matrix $F = [F_{i,j}]$, $i, j \leq 1$ find a

nonzero vector $x \in \mathbf{R}^n$ and a number λ that satisfy the equation $\mathbf{F}x = \lambda x$. Number λ is called the eigenvalue of the matrix \mathbf{F} and x is the associated right eigenvector of \mathbf{F} . Gershgorin's theorem states that each eigenvalue λ of the matrix \mathbf{F} satisfies at least one of the following inequalities

$|\lambda - f_{i,i}| \leq r_i$, where r_i is the sum of all off-diagonal entries in row i of the matrix $|\mathbf{F}|$.

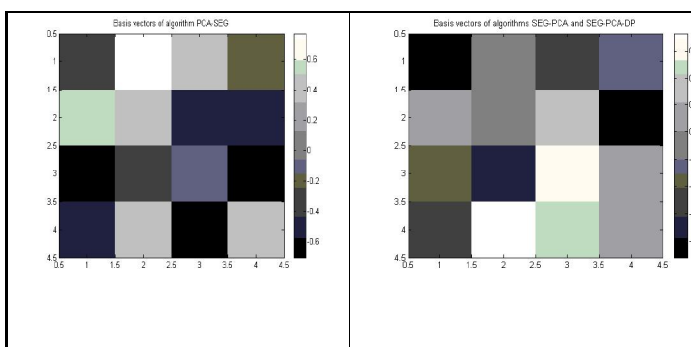
Information about each circle (see table 3) coordinates of the origin and its radius is contained in successive rows of the matrix C in each algorithm PCA process.

Table 3 Gershgorin's circles and Eigenvalue of Matrix F



In all our figures of PCA, the grayscale in each plot is spread linearly from highest negative (white) to highest positive (black) value and so blocks with similar shades correspond to their coefficients having similar values.

Table 3 Basis Vectors of Algorithms.



VI. CONCLUSION

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. This practical application shows how strong the latent structure is in real data. We do the segment similar in some cases really and underling common basis.

REFERENCES

- [1] Bingham Ella , et al , Segmentation and Dimensionality Reduction - HIIT Basic Research Unit , University of Helsinki and Helsinki University of technology Finland August 2005.
- [2] Courrieu, P., Fast Computation of Moore-Penrose Inverse Matrices, Neural Information Processing - Letters and Reviews, Vol.8, No.2, August 2005.
- [3] Dorf, R.C, Modern Control System, Addison –Wesley, Reading, Mass., 1992.
- [4] Föllinger, O., Franke, D., Einführung in die Zustandsbeschreibung dynamischer Systeme. Oldenbourg Verlag, München, 1982
- [5] 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
- [6] 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
- [7] Mikenina,Geb, Angstenberger , Larisa, Dynamic Fuzzy Pattern Recognition, Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online ver,2000