Underlying Structure of the System, Based on the Behavior of the Dynamical System

Hernando Castañeda Marín
Doctorado en Ciencias Aplicadas ULA
Universidad de Pamplona
Facultad de Ingenieria, Pamplona N.S
Colombia
hcastaneda@ula.ve

Wladimir Rodríguez Graterol
Departamento de computación
Universidad de los Andes
Facultad de Ingenieria, Mérida-Edo Mérida,
Venezuela
wladimir@ula.ve

Eliézer Colinas Morles
Departamento de Control y Automatización
Universidad de los Andes
Facultad de Ingenieria, Mérida –Edo Mérida
Venezuela
ecolina@ula.ve

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 problems that until now have found 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 a technique to analyze dynamical systems 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.

Key-Words: - Pattern recognition, Dynamical system, Complex systems, Artificial Intelligence, Similarity.

1 Introduction.

Similarity plays a fundamental role in the theories of knowledge and behavior and has been extensively studied in the literature of psychology.

Traditionally, dynamic systems have been studied using formal mathematical theories.

However, these approaches to system modeling perform poorly for complex, nonlinear, chaotic, and uncertain systems. We believe that a possible way to study and analyze such dynamical systems 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 systems” this approach is motivated by “Case-Based-Reasoning”; this 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 system analysis –e.g. trying to find a mathematical model that describes output as a function of state variable and due to the fact that input perform poorly when dealing with complex systems. This may be due to their nonlinear, time-varying nature or to uncertainty in the available measurement.

We can approach the analysis of dynamic systems in two different ways: the first is based on the existence of a state measuring mechanism in the form of a mathematical model; In the absence of such a measuring mechanism, we must resort to some perceptual mechanism, that allows 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 of the motivations of this dissertation is to discover ways to use structural similarity as mechanics to study dynamic systems.

2.1 Mathematical Description and Modeling of 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 analysis in the domain of 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 systems 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 set of state variables in any instant determines the state of the system at 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 input 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 $\mathbf{X}(t)$ 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 the way to this point in the space of states is known like as trajectory of the system. For an initial state and end state given 1 an infinite number of input vectors exist that correspond to trajectories with start and end points.

On the other hand, through a point on the state space only one trajectory passes. 3 .

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

2.1.1 Classes of Linear Dynamic System.

The state of transition of the dynamic system in the internal space and the mapping from the space of internal states to the space of observations is modeled by the following linear 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_{x_t} $0, Q^{(i)}$ 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 the input zero of the system. In other words, these eigenvalues determine 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 noise process are zero in the equation (1). Using the decomposition of the eigenvalue in the transition matrix the following expressions are obtained:

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

The state in the time t can be resolved with initial conditions x_0 this way:

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

Where λ_p and e_p are the corresponding eigenvalue and eigenvectors and the pondered value α_p is determined from the initial state x_0 in the complex plane.

From this, the general patterns of a system can be categorized by the position of the eigenvalue (poles) in the complex plane. The determination of the oscillatory states are based the following rules.

- It oscillates if at least one eigenvalue is negative or complex.
- It doesn't oscillate if all the eigenvalue are real numbers.

The absolute value of the eigenvalue determines the convergence or divergence state in the way:

"It diverges if at least one value of the eingenvalue is greater that one.
. "It converges if all the absolute values of the eigenvalue are less than one.

Figure 1 shows states of trajectories with two-dimensional states. The systems can generate patterns for temporary variation if and only if this pattern converges to zero. (In control terms is said that the system to stable).

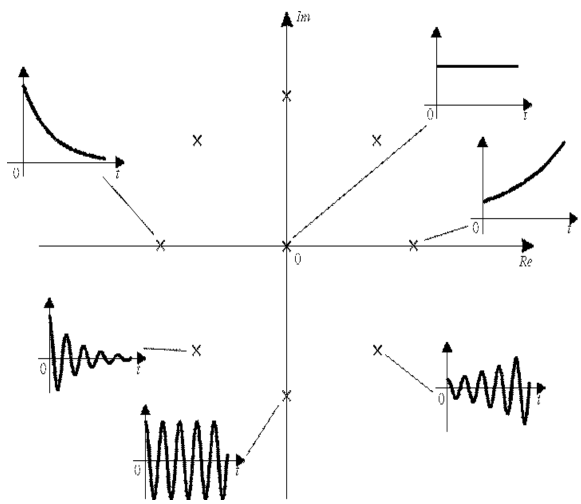

Fig 1 Examples of class of dynamic in the complex plane

2.2 Dynamic Pattern Recognition Base

Let us consider a complex system that assumes different states in the course of time. Each state of the system in the instant of time is considered as an object to classify. If a dynamic system is observed temporarily, the variable value of the features is a dependent function of the time.

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

The objects receive the name of dynamic if they represent measurements or observations of a dynamic system and contain the 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 time. The dependent function of time is represented by a trace, or trajectory, for each object from its initial state to its current state in the space of features.

Figure 2 shows trajectories of objects in a space of two-dimensional features. If the form of the trajectories is chosen as the criterion of similarity the trajectories then three clusters of objects can be distinguished A,C , B,D,E,G y F,H . These three clusters are different to those that are recognized as static objects in an instant in time [6]

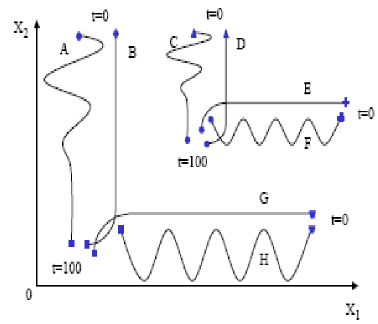

Fig 2 Objects in the two-dimensional space of features

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 and G,H .

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

2.2.1 Similarity Measures Based on the Characteristics of Trajectories.

In the previous section we considered a criterion for the comparison between two trajectories. Two similarity types between trajectories are considered:

1. Pointwise Similarity: the smaller Pointwise distance between two trajectories in feature space. The greater measure of similarity between two trajectories is a criterion.
2. Structural Similarity: the best match of two trajectories in form, evolution, characteristics, and the greatest similarity between two trajectories is criterion.

To determine the structural similarity structural relevant aspects of the behavior of the trajectories are specified depending on a concrete application based on physical properties of the trajectories (e.g. slope, curvature, smoothness, position and value of inflection points) these can be selected, and are then used as comparison criteria.

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

2.2.2 Structural Similarity Based on Slope and Curvature Trajectories.

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

$$cv_k = \ddot{x}_k = \frac{\dot{x}_k - \dot{x}_{k-1}}{t_k - t_{k-1}} \qquad (4)$$

Where $\dot{x}_k$ denotes the coefficient of the first derivative at point x_k and given as:

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

Substituting the previous equation in the equation of the curvature, you arrive at the following equation based on the values of the original trajectories as:

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

The distinctive characteristic when the curvature is considered is the sign of the coefficient of the 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 a negative in a certain period of time, the trajectory is concave (near to the low point). If the coefficient is equal to zero at some point is call the inflection point, not bend is found 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 characteristic qualitative for a description of the segments.

To derive quantitative information starting from the segments, these 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 3a.

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.

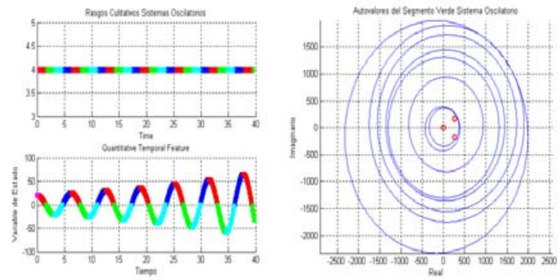

Fig 3 a) Clustering results of an oscillatory system and b) Eigenvalue in complex plane.

3.1 Clusters of Trajectories as Dynamic Systems (Dynamic Objects)

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 $\theta_i \;\; i = 1,\ldots,N$ 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 $y_1^T \triangleq y_1,\ldots,y_T$ 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 $D_1, D_2, \ldots, D_N$ (for example the number of dynamic systems and their parameters $w_1,\ldots,w_N$ for some interval groups I (For example to segment and to label the sequence) from the sample of training. y_1^T where the number of intervals K is also ignored.

We show how the behavior of any states variable in linear system can be decomposed into several modes of behavior, each characterized by an eigenvalue, that is, the time trajectory of state variable i can be expressed as:

$$x_i(t) = w_{i1}m_1(t) + ... + w_{ij}m_j(t) + ... + w_{in}m_n(t) + u_i$$

Where $w_{i,j}$ is a constant term representing the significance of mode j to state variable i, $m_j(t)$ is the value of the j^{th} mode of behavior at time t; and u_i is a constant term. The modes of behavior of a linear system are function of the eigenvalue λ of the jacobiana matrix that characteristic the system [Ogata, 1999]. Where $m_j(t)$ is a function of the form?

$$m_j = \begin{cases} \exp \operatorname{Re}\left[\lambda_j\right]t \quad \text{si} \operatorname{Im}\left[\lambda_j\right] = 0 \\ \exp \operatorname{Re}\left[\lambda_j\right]t \ sen \ \operatorname{Im}\left[\lambda_j\right]t + \theta \ = 0 \text{ else} \end{cases}$$

Behavioral is picture in figure 1.

3.2 System Identification Based 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 eingenvalues 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:

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

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

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

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.

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

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

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

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:

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

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 eingenvalue 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 eingenvalue 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

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:

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

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:

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

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

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

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

We begin by giving some result for locating and bounding the eigenvalues of a matrix $F^{(i)}$ (see figure 3 right side). We give a picture in the complex plane of what the circles form a connected set S, not disjoint, then S contains exactly n of the eigenvalues of $F^{(i)}$, counted according to their multiplicity as roots of the characteristic polynomial of $F^{(i)}$.

4 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.

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