Application of intelligent classification techniques to the TJ-II Thomson Scattering diagnostic

J. Vega, I. Pastor, J. L. Cereceda, A. Pereira, J. Herranz, D. Pérez, M. C. Rodríguez
Asociación EURATOM/CIEMAT para Fusión, Madrid, Spain

G. Farias, S. Dormido-Canto, J. Sánchez, R. Dormido, N. Duro, S. Dormido
Dpto. Informática y Automática - UNED, Madrid, Spain

G. Pajares, M. Santos, J. M. de la Cruz
Dpto. Arquitectura de Computadores y Automática - UCM, Madrid, Spain

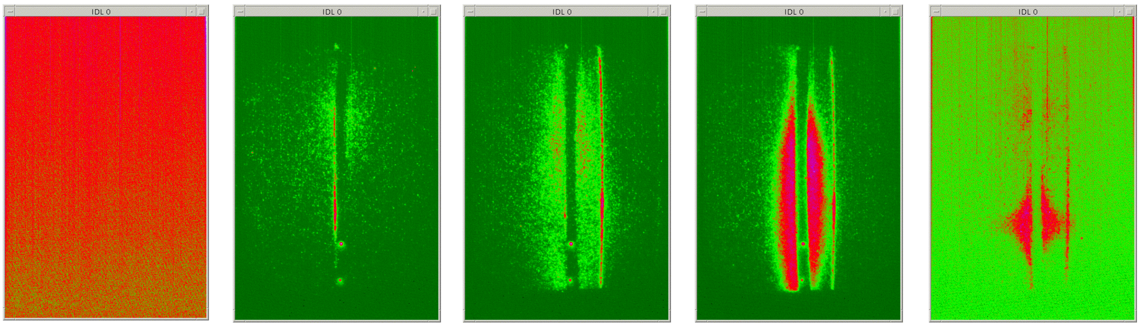
ABSTRACT

Intelligent classification techniques can help to automate data analyses processes in a fusion environment. The data acquisition program of the TJ-II Thomson diagnostic has been synchronized with the TJ-II synchronization system and also, an automated image classification method has been implemented. The classification stage has been based on “Support Vector Machines” mainly due to the efficiency of computation during both system training and image evaluation. After image acquisition, the classifier identifies the kind of image that has been produced and, according to it, the corresponding analysis processes can be executed.

1. INTRODUCTION

The TJ-II Thomson Scattering (TS) diagnostic [1, 2] has been upgraded to increase its automation level as a previous step to work in a unattended manner. On the one hand, the diagnostic control program was integrated into the TJ-II asynchronous event distribution system (AEDS) [3] to synchronize the diagnostic operation with the discharge production. On the other hand, we have developed a specific mechanism to be able to execute automatic analysis processes after a laser shot.

TS images are 2D spectra, with the horizontal and vertical axes displaying respectively scattered wavelength and position along a plasma chord. The type of image captured determines the kind of analysis to perform. TJ-II TS images can be essentially of five different types. These classes correspond respectively to CCD camera background (fig. 1a), measurement of stray light without plasma or in a collapsed discharge (fig. 1b), image during ECH phase (fig. 1c), image during NBI phase (fig. 1d) and image after reaching the cut off density during ECH heating (fig. 1e).



(a) (b) (c) (d) (e)

Fig. 1: Image patterns in the TJ-II Thomson Scattering

2. CLASSIFICATION PROCESS

The aim of objects classification is to find a rule, based on external observations or training elements, that allows assigning each object to anyone of several possible classes. There are two big stages to implement in a classification process: features extraction and objects sorting [4]. The first one consists of performing some pre-processing on the objects trying to extract specific differentiating features. The second stage groups the objects into a set of classes.

In the TJ-II Thomson Scattering case, the objects to classify are images. Each image has (385x576) pixels, *i.e.* 221760 possible attributes. Two different set of features were considered. Firstly, we took into consideration a general classification scheme based on a wavelet transform (WT). Secondly, we tried to take advantage on some physical knowledge about the images and we tested several sets of reduced characteristics. Support Vector Machine (SVM) has been used for pattern recognition.

2.1 WAVELET TRANSFORM

Analysis of bi-dimensional signals is getting great improvements by using Wavelet based methods. Due to the fact that the WT decomposition is multi-scale, images can be characterized for a set of approximation coefficients and three sets of detailed coefficients (horizontal, vertical and diagonals). The approximation coefficients represent coarse image information (they contain the most part of the image's energy), whereas the details are close to zero, but the information they represent can be relevant in a particular context.

We have found that the best coefficient to characterize the TJ-II Thomson images is the vertical detail, when selected the Haar Wavelet at level 4 [5]. With these setting, the attributes are reduced from 221760 to 900 (0.4% of the initial attributes).

2.2 REDUCED CHARACTERISTICS

TJ-II TS images follow a general pattern shown in figure 2. We defined different sets of features based on statistical properties of the images and relevant bands: sum(pixel intensities), standard deviation (intensities), intensity profiles of rows, intensity profiles of columns, intensity centre of gravity (by rows or columns), background subtraction and so on. Each set has approximately 10 attributes (0.005% of the initial characteristics).

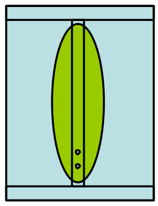


Fig. 2: General pattern of the TJ-II Thomson Scattering images

2.3 SUPPORT VECTOR MACHINES

SVM is a very effective method based on kernels for general purpose pattern recognition [4, 6, 7]. In a few words, given a set of input vectors which belong to two different classes, the SVM maps the inputs into a high-dimensional feature space through some non-linear mapping (kernel functions), where an optimal separating hyper-plane is constructed in order to minimize the risk of misclassification. The hyper-plane is determined by a subset of points of the two classes, named support vectors.

3. THOMSON SCATTERING DIAGNOSTIC UPGRADE

To incorporate both synchronization with the TJ-II operation and image pattern recognition, we have developed several software applications. First, the set of functions related to image feature extraction, learning system and pattern recognition. All of them were programmed in MATLAB. Second, a server program is in execution on a Windows 2000 computer that is in charge of receiving images (not only during TJ-II discharge production) and returning the classification. Moreover, the program maintains a relational database with the classification results. It is a C++ program based on BSD sockets for network communications that invokes the corresponding MATLAB functions to perform the classification process. Third, we have several C programs for the TS control workstation. On the one hand, the program that controls the diagnostic during TJ-II operation and, on the other hand, a set of tools for off-line interaction with the classifier system: off-line image sorting and debugging wrong classifications. Figure 3 shows the analysis automation during operation.

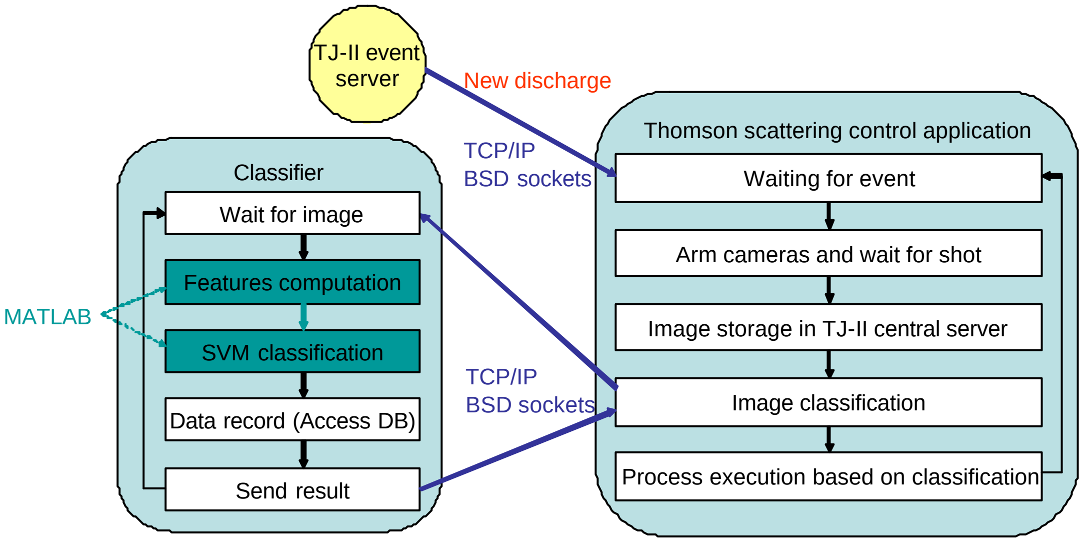


Fig. 3: Analysis automation

4. RESULTS

The WT+SVM technique is very robust and success rate was 92.7% in the last TJ-II experimental campaign (over a 98% in earlier campaigns). The classifier did not know to assign a class in a 5.5% of the cases and the rate of wrong classifications was 1.8%. SVM allows us to improve our system by introducing other kernels. Success rate with reduced characteristics and SVM was 50% (95% in preceding campaigns). However, this technique is less robust than the previous one and the learning process is extraordinarily dependent on the set of training images.

REFERENCES

[1] C. J. Barth et al. Review of Scientific Instruments. 70 (1999) 763

[2] J. Herranz et al. Fusion Engineering and Design. 65 (2003) 525

[3] J. Vega et al. Fusion Engineering and Design. 71 (2004) 117

[4] V. Vapnik. “Statistical Learning Theory”. John Wiley & Sons, INC, (1998)

[5] G. Farias et al. XXV Jornadas de Automática. Ciudad Real (Spain). ISBN 84-688-7460-4. (2004)

[6] B. Schölkopf and A. J. Smola. “Learning with kernels”. MIT Press. (2002)

[7] R. O. Duda, P. E. Hart and D. G. Stork. “Pattern Classification”. 2nd edition. John Wiley (2001)