A. Murari, J. Vega, D. Mazon, J.F. Delmond, M. Gelfusa, A. Szappanos
and JET EFDA contributors

Image Manipulation for High Temperature Plasmas

Image Manipulation for High Temperature Plasmas

A. Murari[1], J. Vega[2], D. Mazon[3], J.F. Delmond[4], M.Gelfusa[5], A.Szappanos[6]
and JET EFDA contributors*

JET-EFDA, Culham Science Centre, OX14 3DB, Abingdon, UK

[1]*Consorzio RFX-Associazione EURATOM ENEA per la Fusione, I-35127 Padova, Italy*
[2]*Asociación EURATOM/CIEMAT para Fusión, Avenida Complutense, 22, 28040 Madrid, Spain*
[3]*EURATOM-CEA, CEA Cadarache, 13108 Saint-Paul-lez-Durance, France*
[4]*Arts et Métiers ParisTech Engineering College (ENSAM) 75013 Paris.*
[5]*Associazione EURATOM-ENEA - University of Rome"Tor Vergata" , Roma, Italy*
[6]*Association EURATOM-HAS*
* See annex of F. Romanelli et al, "Overview of JET Results" ,
(Proc. 22 [nd] IAEA Fusion Energy Conference, Geneva, Switzerland (2008)).

ABSTRACT

In magnetic confinement Nuclear Fusion video cameras have become routine diagnostics, which can produce gigabytes of data per discharge. New tools and methods are required to manipulate the frames of these videos to obtain the required information. New algorithms have been developed, which implement pattern recognition methods to search the repositories of videos on the basis of their content and not on their address. Real time automatic analysis of videos has motivated the development of machine learning tools to classify objects in images. The identification and tracking of objects in the videos of Tokamak diagnostics requires the computation of image descriptor, which are insensitive to rotation, translation and rescaling. The Hu moments have proved to be quite effective in performing this image analysis task.

1. INTRODUCTION

Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. Particularly widespread has become the manipulation of bidimensional signals. In the last decades, digital video cameras have improved enormously and they are now found in a variety of devices, ranging from cellular phones to portable digital assistants and game consoles. The easy capture and storage of images, together with the intuitive appeal of visual information, has motivated the development of new methods and techniques to analyse these bidimensional signals. Indeed, even if the amount of information contained in videos is enormous, very often it is not in a format that automatic systems can easily profit from. Moreover, the amount of frames to analyse becomes very soon so large that manual analysis is prohibitive. Therefore the extraction of the required information from large repositories of videos has become a quite important field of research in computational science and artificial intelligence. The goals of this manipulation of images to extract useful information can be divided into three main categories [1]:

* Image Processing image in -> image out
* Image Analysis image in -> measurements out
* Image Understanding image in -> high-level description out

In the first case, the input to the manipulating tool is an image and the output is again an image. The term Image Analysis is reserved to the cases in which frames have to be processed to obtain some form of quantitative measurements. In the case of Image Understanding the objective consists of deriving some form of high level interpretation of the images.

The quality of video cameras is such that nowadays they are commonly used also in many field of research, including Magnetic Confinement Nuclear Fusion (MCNF). Conceptually the tasks and objectives of image manipulation for scientific purposes are the same previously mentioned but often with some particular specificities, which require dedicated developments. In the field of MCNF the main difficulties presented by image manipulation can be grouped in three main categories:

1) Information retrieval
2) Extraction of information in real time
3) Derivation of quantitative physical information

In MCNF cameras can produce very large amounts of information for every experiment. In JET [2], the largest MCNF experiment in the world, the wide angle fast camera can operate to a maximum of 250 Kframes/second and produces gigabytes of data per discharge (in JET a typical discharge can last for a few ten of seconds and 30 or 40 discharges per day are routinely performed during the experimental campaigns). Retrieving the required information in the resulting enormous image reservoir is a significant task, which has motivated the implementation of new methods based on structural pattern recognition (see section 2). A different category of issues is presented by the needs of real time control. Progressively video cameras, both in the IR and visible range of wavelengths, are being used for control of the experiments and even for machine protection. The need to process the frames of the cameras quickly and reliably is sometimes a challenge, which requires the deployment of quite sophisticated machine learning tools and image descriptors (see section 3). Summary and possible lines of future research are the subjects of the last section of the paper.

The videos, which have been used to obtain the results shown in this paper, have been collected by one of JET most interesting viewing system for image processing applications. It hosts cameras installed on a dedicated endoscope providing a wide-angle view (field of view of 70 degrees) in the infrared range (3.5 to 5µm) and in the visible. The wide angle view of the system includes the main chamber and the divertor [3]. The diagnostic consists of an endoscope formed by a tube holding the front head mirrors, a Cassegrain telescope, and a relay group of lenses, connected to the camera body. To increase the reactor-relevance of the project, mainly reflective optical components have been installed, since they can better cope with high neutron radiation. The fast visible camera located on this endoscope has a 1024×1024 CMOS pixel detector, which can be acquired full frame up to 3kHz. The maximum frame rate of the camera is 250kframes/s (for a reduced frame of 128×16 pixels).

2. IMAGE PROCESSING: PATTERN RECOGNITION FOR INFORMATION RETRIEVAL

Traditionally computer databases are address based: the location in memory must be known to retrieve the required information. On the contrary human memories are content based, i.e. somehow the information is retrieved on the basis of its content. The main idea behind content based databases is represented in figure 1. From a practical point of view, in the case of traditional databases the input to be provided is the address and then the libraries retrieve the corresponding samples. On the other hand, pattern recognition libraries need only an example of the required pattern and then they search the entire database to find the shots and time intervals where the same or similar patterns are present. The approach can be used to signals of any dimensionality and therefore also to images. Content based databases would be very useful in science. In MCNF in particular, the first approach to the data is certainly based on the content of the signals or the images acquired during an experiment. With traditional databases this requires in principle to scan the whole database until the required

information content (typically a signal of a specific form) is found. As mentioned, the cameras located on JET wide angle endoscope have the potential to produce gigabytes of data per JET discharge (30-40 discharges per day are normally performed). Exhaustive manual searches of hundreds of thousands of frames per discharge or per day are not an option and therefore automatic techniques are required. To implement the concept of content based information retrieval the starting point is the observation that the first step in data analysis for fusion typically consists of a visual screening of the signals or the images. Since this is a very high level visual process, the methods developed to help the physicists in this respect are based on structural pattern recognition. This approach hinges on a concept of similarity between images, which is based on the structure of the image and not on involved mathematical concepts, such as correlation or covariance, which are difficult to appreciate by simple visual inspection. Indeed the main objective of the proposed method is to help the user during the first screening phase of the data.

In more detail, the Haar transform is first applied to the images. Then the coefficients of the transform are grouped in different classes depending on their value. To each of these classes a letter is associated, transforming a matrix of pixels (the original image) into a bidimensional array of letters. The obtained matrix of letters can be codified using a relational database (PostgreSQL for the results presented in this paper). The patterns to be retrieved are bidimensional. The strategy of the search consists of looking for the first rows in the pattern. Once a match for this first row has been found, it is then determined whether also the other rows of the bidimensional pattern coincide. The implementation of this search by mature relational database technology allows performing this search quite efficiently (even if increased speed is desirable as mentioned later).

An example of the results is presented in figure 2, which shows the user interface available to the users at JET. A particular region of interest of a visible camera image, basically the divertor, has been selected (part within the square in the central frame of figure 2) and then the pattern recognition routines manage to identify the frames in the database where the same or a similar pattern is present (as shown in the right hand part of the figure). For a database of 26000 frames and a total of more than 6 Gigabytes of data, the additional space in memory is less than 0.5% of the original images and is therefore negligible. The search takes a matter of minutes and therefore, even if of course this is orders of magnitudes better than a manual search, parallelisation techniques are being developed to improve this aspect.

3. IMAGE ANALYSIS: DETECTION AND TRACKING OF OBJECTS IN REAL TIME

In the last decade or so the measurements of infrared cameras have been systematically used for real time protection. In particular various algorithmic solutions have already been found to detect hot spots, regions of unusual high temperature on the internal wall of the vacuum vessel, in real time. More sophisticated approaches, including machine learning tools such as Neural Networks and Support Vector Machines, are proving indispensable to analyse the more complex images detected by visible cameras.

Reliable real time detection of instabilities, which leave a clear signature in the videos of visible

cameras, such as MARFEs [4] (Multifaceted Asymmetric Radiation From the Edge) and ELMs (Edge Localised Modes) [2] could provide important information to improve the safe operation of the devices and to increase the success rate of disruption prediction techniques. Identification of UFOs (unidentified objects which can appear in the frames of JET videos) would also contribute significantly the safe operation of large devices such as JET. Unfortunately the identification and tracking of these phenomena in JET visible cameras has proved to be quite challenging. To be useful in real time, the image manipulation algorithms have of course to be fast enough to allow enough margins of intervention. On the other hand, they have also to provide reliable information, and this can be problematic given the nature of the movies of video cameras in MCNF. Some of the phenomena to be followed are quite rapid and therefore their evolution can be difficult to follow even at the high frame rate of JET fast camera. Moreover the content of the individual frames can be very complex, difficult to decipher and tends to vary significantly not only from discharge to discharge but sometimes even during different phases of the same shot.

An indicator of the complexity of these images is provided by the challenges posed by even the most elementary and typically preliminary step of thresholding, which the typical first operation adopted to isolate the objects of interest from the rest of the image. In the case of visible videos it is typically impossible to perform a constant thresholding because the luminosity of the frames changes too much and the background can be very different from one frame to the next. Therefore the only general approach found to work properly consists of first blurring the image (by substituting the grey level of each pixel with the average of a suitable number of neighbours). Then the blurred frame is subtracted from the original one and the threshold is applied to the difference. This example gives an idea of the difficulties involved in image manipulation and in particular image analysis for fusion.

More generally the identification and tracking of objects in JET videos requires quite sophisticated methods, which are illustrated by the example of detection of MARFEs. MARFEs are instabilities, which occur at high density and can not only reduce confinement but also cause disruptions, sudden losses of plasma confinement abruptly terminating a discharge and even putting at risk the integrity of the devices. An example of the bright ring of radiation as seen by JET fast camera is shown in figure 3. As can be seen from the sequence of frames, the objects to be detected change shape and position quite rapidly in time. Moreover, inside the plasma there are other phenomena, typically instabilities, which leave similar signatures on the videos of JET visible cameras. Identifying MARFEs therefore requires a series of quite advanced indicators, which can properly describe the geometry of the radiation pattern left by the instability in the frames of the visible camera. In terms of descriptors, the Statistical Moments called Hu moments have proved to be quite appropriate to capture the man characteristics of MARFEs (see later). These indicators have then been given as input to an SVM classifier [5] to perform the real time analysis, because algorithmic solutions tend not to generalise well enough.

The object tracking tool, which has been developed to identify MARFEs, is based on the evaluation of Statistical Moments for binary image analysis. As previously mentioned, the background subtraction step, no matter how complicated, gives as output a binary image. One of the advantages

4

of binary images is that they can be easily processed and analyzed. In particular the geometrical properties of the objects can be determined by means of statistical moments. Since the objects to be detected can be located in different parts of the frames and can have different orientations, it is essential to described them with indicators that are invariant to translation, rotation and scaling. The use of the statistical moments with these properties is crucial for this application, in which objects, located in very different positions inside the field of view, have to be identified.

In general, given a bidimensional function f(x,y), it is possible to define a moment of grade p+q as follows:

$$M_{pq} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^p y^q f(x,y)\, dx \cdot dy \tag{1}$$

For a binary image, whose pixels have intensity indicated by the function I, a discrete form of equation (3) can be written as:

$$m_{pq} = \sum_x \sum_y x^p y^q I(x,y) \tag{2}$$

Two additional parameters, which are very important in binary image processing, are the centres of mass in the x and y directions:

$$\bar{x} = \frac{m_{10}}{m_{00}}; \qquad \bar{y} = \frac{m_{01}}{m_{00}} \tag{3}$$

Using the centres of mass, the centralized normalized moments can be defined as mpq:

$$_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q I(x,y) \tag{4}$$

Combining suitably various central moments, Hu manage to define [6] a set of 7 moments, which are translation, scale and rotation invariant; their mathematical expressions are reported in relations (5)-(11).

$$\varphi_1 = \eta_{20} + \eta_{02} \tag{5}$$

$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + (4\eta_{11})^2 \tag{6}$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{7}$$

$$\varphi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{8}$$

$$\varphi_5 = (\eta_{30} - 3\eta_{12}) \cdot (\eta_{30} + \eta_{12}) \cdot \left[(\eta_{30} + \eta_{12})^2 - 3 \cdot (\eta_{21} + \eta_{03})^2 \right] +$$
$$+ (3\eta_{21} - \eta_{03}) \cdot (\eta_{21} + \eta_{03}) \cdot \left[3(\eta_{30} + \eta_{12}{}^2) - (\eta_{21} + \eta_{03})^2 \right] \tag{9}$$

$$\varphi_6 = (\eta_{20} - \eta_{02}) \cdot (\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] +$$
$$+ 4\eta_{11}(\eta_{30} + \eta_{12}) \cdot (\eta_{21} + \eta_{03}) \tag{10}$$

$$\varphi_7 = (3\eta_{21} - \eta_{03}) \cdot (\eta_{30} + \eta_{12}) \cdot \left[(\eta_{30} + \eta_{12})^2 - 3 \cdot (\eta_{21} + \eta_{03})^2 \right] +$$
$$+ (\eta_{30} - 3\eta_{12}) \cdot (\eta_{21} + \eta_{03}) \cdot \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \tag{11}$$

An example of the invariant character of the Hu moments is given in figure 4. On the top row a view of JET divertor captured by JET fast visible camera is shown rotated of 45 degree from 0 to 360. It is easily to see how the Hu moments do not change significantly when the object is rotated. The same applies to rescaling and translation.

The centers of mass and the Hu moments have been given as input to an SVM classifier to discriminate between the signature of MARFEs and other objects in the frames of JET visible camera. The radial basis function is the kernel which has provided the best results. A database of 1300 frames has been prepared for this study. All the frames in the database have been analysed manually to determine the nature of the objects present in them. As usual about 60% of the database has been used for training and the rest for testing. The best results in terms of classification accuracy has been found using the centres of mass and the first Hu moment as input to the SVM classifier. The performance of the classification tool are reported in figure 5, which shows the quite positive success rate achieved, particularly taking into account that very difficult frames have been included in the database. In the case of mass centres and the first Hu moment as inputs, the success rate is in excess of 98%. A visual example of the bright emitting region generated by a MARFE and properly identified by the classifier is shown in figure 6.

FUTURE PROSPECTS

Both the information retrieval tools based on pattern recognition and the real time classifiers for tracking pose some problems in terms of computational time. In the case of the pattern recognition algorithms, the main line of future research is parallelization. In the case of real time applications, in addition to reliability and speed, predictability is also an important requirement. It is important that the computational time does not vary significantly with the content of the images. To meet this requirement the adoption of morphological operator and even the computational paradigm of Cellular Nonlinear Networks are certainly very promising alternatives.

REFERENCES

[1]. M. Sonka, V.Hlavac, R.Boyle "Image processing, analysis and machine vision" Thomson, London, 2008 Third edition.

[2]. J. Wesson "Tokamaks" Clarendon Press-Oxford 2004

[3]. E. Gauthier et al. Fusion Engineering and Design, Volume **82**, Issues 5-14, October 2007, Pages 1335-1340.

[4]. B. Lipschultz, J. Nucl. Mater. **145-147** (1987) 15.

[5]. C.M. Bishop "Pattern recognition and Machine Learning" Springer, Singapore 2009.

[6]. M. K. Hu, IRE Trans. Info. Theory, vol. IT-8, pp.**179–187**, 1962
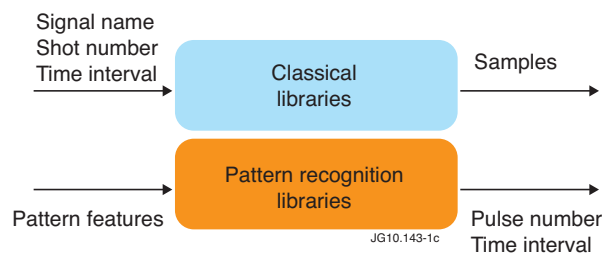
Figure 1: Different approaches to information retrieval of pattern recognition libraries compared to traditional address based databases.
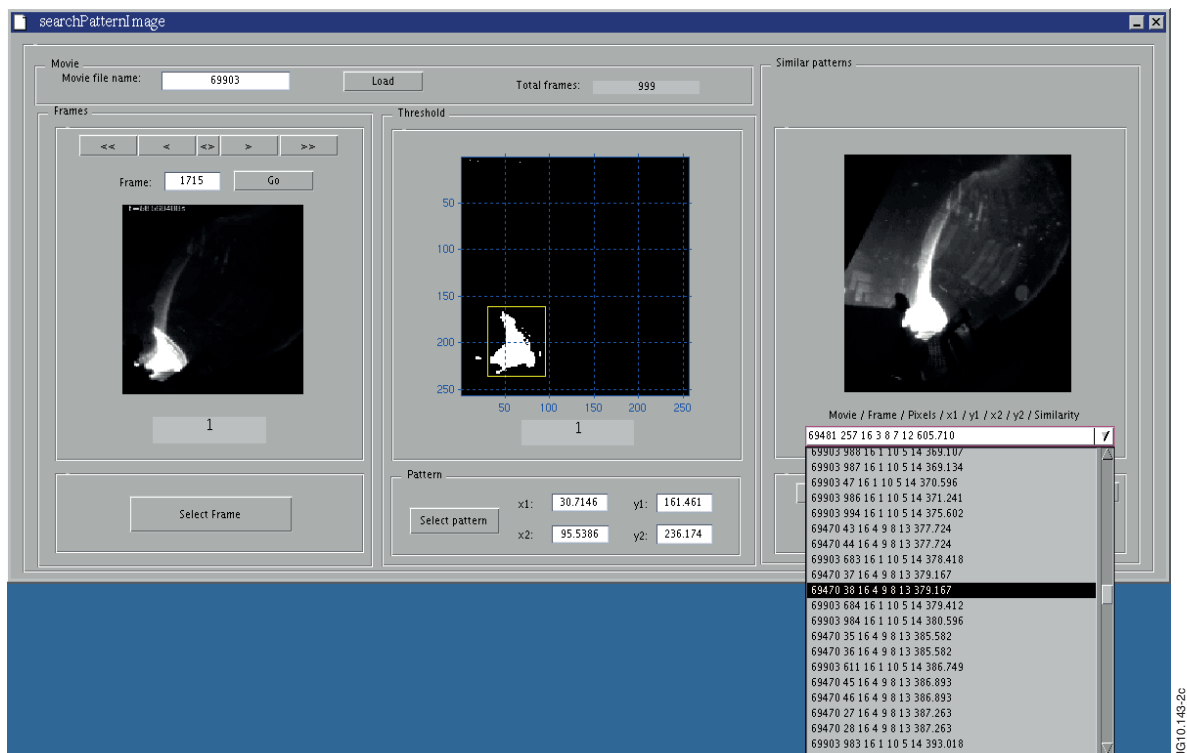

Figure 2: An example of retrieval of patterns within images using the structural pattern recognition approach described in the paper. The object to search for is selected with the cursor by the user (central image) in a generic frame. The pattern recognition algorithms identify the frames in the database, which contain the same or similar objects. An example of these frames with similar patterns and their list is shown on the right part of the figure.
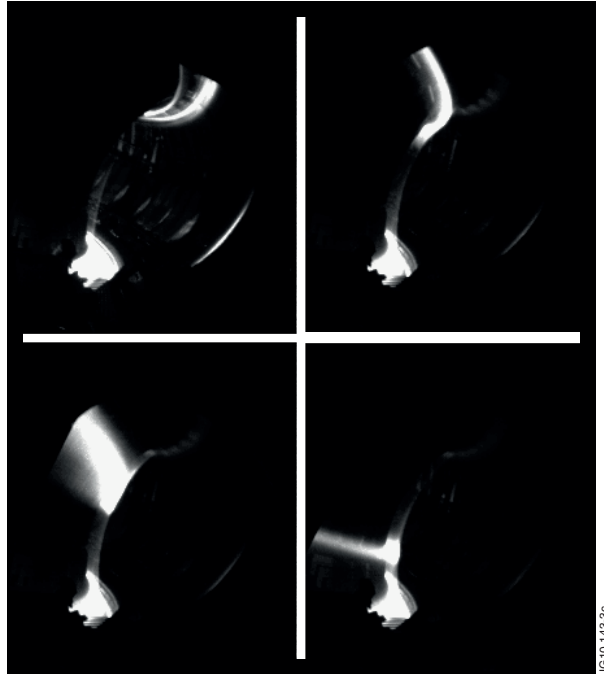
Figure 3: Sequence of frames taken by JET fast visible camera showing the MARFE evolution along the inner wall of JET vacuum vessel.

Insensitivity to rotation

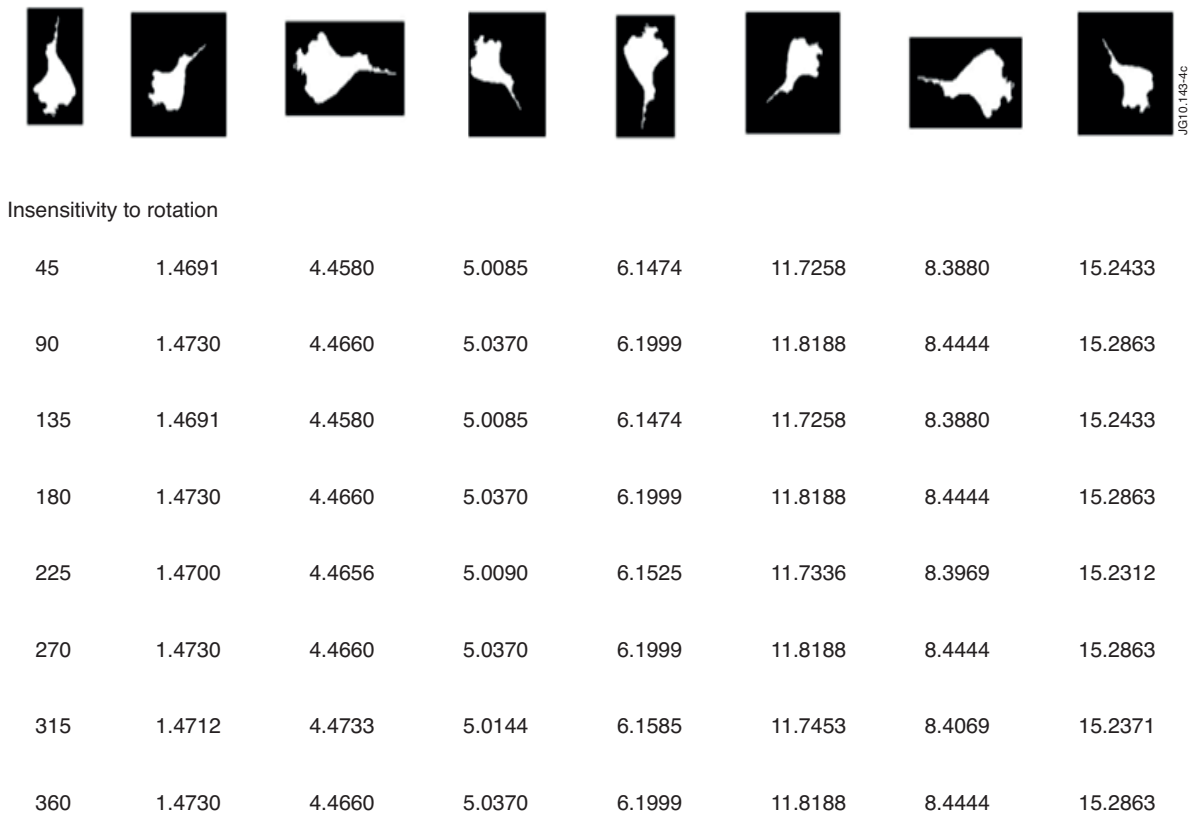|							
45	1.4691	4.4580	5.0085	6.1474	11.7258	8.3880	15.2433
90	1.4730	4.4660	5.0370	6.1999	11.8188	8.4444	15.2863
135	1.4691	4.4580	5.0085	6.1474	11.7258	8.3880	15.2433
180	1.4730	4.4660	5.0370	6.1999	11.8188	8.4444	15.2863
225	1.4700	4.4656	5.0090	6.1525	11.7336	8.3969	15.2312
270	1.4730	4.4660	5.0370	6.1999	11.8188	8.4444	15.2863
315	1.4712	4.4733	5.0144	6.1585	11.7453	8.4069	15.2371
360	1.4730	4.4660	5.0370	6.1999	11.8188	8.4444	15.2863

Figure 4: The top row shows a detail of JET divertor as measured by JET fast visible camera rotated of 45 degrees from 0 to 360 degrees. The 7 Hu moments are reported for the various angles of rotation in the respective rows.
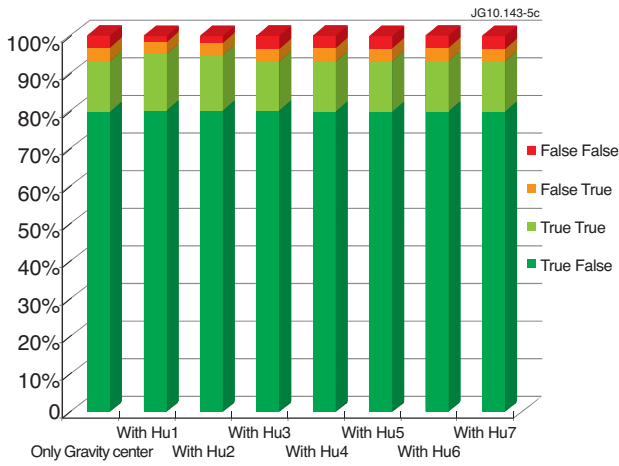
Figure 5: Success rate of the SVM classifier using various inputs. In each column the indicated Hu moment is added to the previous one as input to the classifier.

Figure 6: Example of correct detection of a MARFE in a frame of JET fast camera.