EFDA–JET–CP(14)05/19

T. Craciunescu, A. Murari, M. Gelfusa, I. Tiseanu, V. Zoita,
G. Arnoux and JET EFDA contributors

Advanced Methods for Image Registration Applied to JET Videos

Advanced Methods for Image Registration Applied to JET Videos

T. Craciunescu[1], A. Murari[2], M. Gelfusa[3], I. Tiseanu[1], V. Zoita[1], G. Arnoux[4] and JET EFDA contributors*

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

[1]EURATOM-MEdC Association, NILPRP, Bucharest, Romania
[2]Consorzio RFX, Associazione EURATOM-ENEA per la Fusione, Padova, Italy
[3]Associazione EURATOM-ENEA - University of Rome "Tor Vergata", Roma, Italy.
[4]EURATOM/CCFE Fusion Association, Culham Science Centre, Abingdon, Oxon, UK
** See annex of F. Romanelli et al, "Overview of JET Results", (24th IAEA Fusion Energy Conference, San Diego, USA (2012)).*

Preprint of Paper to be submitted for publication in Proceedings of the 28th Symposium on Fusion Technology SOFT, San Sebastian, Spain
29th September 2014 – 3rd October 2014

ABSTRACT

The last years have witnessed on JET a significant increase in the use of digital cameras, which are routinely applied for imaging in the IR and visible spectral regions. One of the main technical difficulties, in interpreting the data of camera based diagnostics, is the presence of movements of the field of view. Small movements occur due to machine shaking during normal pulses while large ones may arise during disruptions. Some cameras show a correlation of image movement with change of magnetic field strength. For deriving unaltered information from the videos and for allowing correct interpretation, an image registration method, based on highly distinctive Scale Invariant Feature Transform (SIFT) descriptors and on the Coherent Point Drift (CPD) points set registration technique, has been developed. The algorithm incorporates a complex procedure for rejecting outliers. The method has been applied for vibrations correction to videos collected by the JET wide angle infrared camera and for the correction of spurious rotations in the case of the JET fast visible camera (which is equipped with an image intensifier). The method has proved to be able to deal with the images provided by this camera frequently characterized by low contrast and a high level of blurring and noise.

1. INTRODUCTION

During the last decade, digital cameras have undergone rapid development and the growing amount and diversity of obtained images has justified their increasing use in tokamaks for the surveying and control of the experiments and also for retrieving physical information. On JET tens of cameras are now routinely used for imaging on JET in the IR and visible spectral regions. Throughout the image acquisition process different potential sources of movements in the camera and/or their optics may occur leading to translation and rotation of the field of view. The different potential sources of movements in video frames are related to the operation of the devices (e.g. the rapid variations in the magnetic fields) or to plasma phenomena (e.g. instabilities such ELMs and disruptions). Fortunately they typically generate only translations or rotations of the field of view and do not introduce distortions in the camera images. However, the changes in the background illumination, the noise and the blurring effects restrict strongly the class of image registration methods that can be applied for correcting these artifacts.

The process of alignment of two image scenes (the reference and sensed image) taken at different times, from different viewpoints, and/or by different sensors is called image registration. The huge diversity of images to be registered, originating from a wide range of applications such as e.g. remote sensing, medicine, computer vision, earth sciences and the various types of possible degradations which may occur, has led to the development of a wide range of methods which frequently take into account application-dependent characteristics. For a comprehensive survey the reader may refer to the works published by Brown [1] and Zitova et al. [2].

On JET, a series of so called area based approaches have been developed and applied. This class of methods is based on the use of windows with predefined size for estimating the correspondence

between images. Using a correlation coefficient, a measure of similarity can be computed for window pairs from the sensed and reference images and its maximum is searched. Mutual information, which is a measure of statistical dependency between two data sets, can be also an appropriate choice. The registration is achieved by maximization of the mutual information which can be expressed by means of the informational entropy. This class of methods achieved remarkable results in multimodal registration. This a difficult task where images generated by means of very different physical principles must be compared (e.g. anatomical and functional images of the patient's body in medical imaging). On JET the potential of various information theoretic indicators (correlations, entropies and mutual information) for vibrations detection and image registering has been also investigated in a recent study [3].

The Fourier representation of the images (or parts of the images selected by the pre-defined windows) in the frequency domain can be exploited by means of the Fourier Shift Theorem [4]. The cross-power spectrum of the sensed and reference images is calculated and then location of the peak in its inverse is searched for determining the image translation. This class of methods has the advantage of a strong robustness against noise and varying illumination disturbances. The computational time is also a significant benefit for large images. On JET a method developed initially at Tore Supra [5] has been adapted and applied for the vibration corrections of the wide angle view IR camera (KL-7) system which has been installed for the ITER-like wall campaigns [6-7]. In principle this method can achieve real-time capabilities by using field programmable gate array (FPGA) as already proved for the cameras installed on Tore Supra.

The present technique is a feature-based image registration approach. The class of methods starts with a feature detection step. Distinctive features are automatically detected and represented by feature descriptors. The correspondence between the features detected in the sensed image and those detected in the reference image is established by means of a certain similarity measure along with spatial relationships among the features. The developed method aims at the rotation correction of the images provided by the wide angle view fast visible camera (KL-8) installed in JET [8]. The camera is viewing the full poloidal cross-section of the vacuum vessel and is covering a toroidal extent of ~ 90^{o}. The wide angle view is appropriate for the study of pellet ablation, large scale instabilities and plasma wall interactions. Recently the camera has been upgraded with an image intensifier for light amplification and a filter wheel in order to allow the system for spectroscopic plasma imaging filtering atomic lines [9]. The image intensifier is sensitive to magnetic fields so that above a magnetic threshold, image rotation and translation appears. Image processing algorithms for correcting these artifacts have to be implemented. This is a challenging task due to the intrinsic characteristics of the images provided by the camera (noise, blurring, saturation).

2. THE SIFT METHOD

The rotation correction method is based on image registration using Scale Invariant Feature Transform (SIFT) descriptors [10]. These descriptors are inspired and shares a number of properties

with the responses of neurons in inferior temporal cortex in primate vision. Neurons respond to a gradient at a particular orientation and spatial frequency, but the location of the gradient on the retina is allowed to shift over a small receptive field rather than being precisely localized. This allows for matching and recognition of 3D objects from a range of viewpoints [10]. The SIFT algorithm is inspired by this idea, but allows for positional shift using a different computational mechanism.

The potential points of interest in the image are searched over all scales and image locations. To efficiently detect stable key-point locations, the difference of the Gaussian-blurred images at scales and is calculated:

$$DoG(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y,\sigma)) * Im(x, y) \quad (1)$$

where is the image and is a Gaussian of variance σ. *DoG* represents a close approximation of the scale-normalized Laplacian of Gaussian, $\sigma^2\nabla^2 G$ [11] which produces the most stable image features compared to a range of other possible image functions, such as the gradient, Hessian, or Harris corner function [12]. DoG has the advantage of fast computation. The image gradient magnitudes and orientations are sampled around the key-point location. SIFT interest points are identified as local maxima/minima of the DoG images across scales. After scale-space extrema detection and key-point localization, an orientation is assigned to each key-point location based on local image gradient directions. A gradient orientation histogram is computed in the neighborhood of the each key-point. Peaks in the orientation histogram correspond to dominant directions of local gradients. The highest peak in the histogram is detected, and then any other local peak that is within 80% of the highest peak is used to create an additional key-point with that orientation. The points with multiple orientations represent a minority but, according to Lowe (see again Ref. 10), they prove to contribute significantly to the stability of matching between images.

After assigning an image location, scale, and orientation to each key-point, a highly distinctive descriptor for the local image region is constructed by sampling the image gradient magnitudes and orientations around the key-point location. A Gaussian window weighting procedure is used in order to avoid sudden changes in the descriptor. Orientation histograms are constructed summarizing the contents over $n \times n$ subregions. In order to achieve orientation invariance, the coordinates of the descriptor and the gradient orientations are rotated relative to the key-point orientation. The descriptor is formed from a vector containing the values of all the orientation histogram entries. Two images are finally registered by individual key-point matching using an Euclidian distance. The search is performed for key-points at the same scale. The key-points are used as inputs to a nearest-neighbor indexing type method that identifies candidate object matches.

3. APPLICATION FOR THE IR KL-7 CAMERA

The key-point descriptors are highly distinctive. However, in a cluttered image, several key-points may give rise to false matches. A simple approach for removing the false matches has

been implemented by using the so called '3D phase space method' (3DPS) [13]. This method, developed for spike removal, uses the concept of a 3D Poincaré map. The variable, its first and second derivatives, are plotted against each other. The points located outside of the ellipsoid in the Poincaré map are excluded and the method iterates until the number of detected spikes becomes zero. The differences between the (x, y) locations of the possible pair-key-points are calculated together with their average values. The pairs for which the location differences significantly deviate from the average values are removed.

This technique proved to be efficient for the KL-7 camera which provides 'rich content' images and therefore the SIFT method generates a large number of descriptors that densely cover the image over the full range of scales and locations (Figure 1). An efficient vibration correction, with subpixel resolution, has been achieved, as illustrated in Figure 2. and in the supplementary video data accompanying the on-line version of this paper {LINK}.

4. APPLICATION FOR THE FAST VISIBLE KL-8 CAMERA

The number of SIFT key-point matches decreases abruptly in case of the KL-8 camera (Figure 3). Moreover for a significant number of cases, due to image blurring, noise, etc., the matched SIFT descriptors provide an ambiguous information for rotation angle and shift parameters (see again Figure 3).

In order to deal with these equivocal cases, a more flexible matching method must be used. A good candidate is the Coherent Point Drift (CPD) method [14], which is a non-rigid point set registration method. The basic idea of the method is to fit the Gaussian mixture model (GMM) centroids (representing the first point set) to the data (representing the second point set X) by maximizing the likelihood.

$$p(x) = \sum_{m=1}^{M} P(m)p(x|m) \tag{2}$$

where: $p(x|m) = \frac{1}{(2\pi\sigma^2)^{D/2}} exp\left(-\frac{\|x-y_m\|^2}{2\sigma^2}\right)$, is the dimension of the points set, N, M are the number of points in the points sets, $X_{N\times D}$ and $Y_{M\times D}$ are the two point sets. An additional uniform distribution $p(x|M+1) = 1/N$ can be introduced to account for noise and outliers with a weight w ($0 \le w \le 1$). The present implementation uses isotropic covariances and equal membership probabilities $p(m) = 1/M$ for all GMM components.

The correspondence probability between two points and y_m and x_n is defined as the posterior probability of the GMM centroid given the data point: $P(m|x_n) = P(m)p(x_n|m)/p(x_n)$. During the minimization, the GMM centroids are forced to move coherently, as a group, to preserve the topological structure of the point sets. This constraint is introduced by means of the prior:

$$P(y|\lambda) = exp\left(-\frac{\lambda}{2}\Phi(Y)\right) \tag{3}$$

where $\Phi(Y)$ is a regularizing function which enforces smooth motion. Using Bayes theorem, Y is found by minimizing the posteriori probability:

$$E(\theta,\sigma^2)=-\sum_{n=1}^{N} log \sum_{m=1}^{M} exp\left(-\frac{1}{2}\left\|\frac{x_n-y_m}{\sigma}\right\|+\frac{\lambda}{2}\Phi(Y)\right) \quad (4)$$

A smooth motion means a smooth velocity field with no oscillatory behavior. Therefore the regularizing function should provides less energy at high frequency.
Consequently an adequate choice will be $\Phi(Y)\sim\int\frac{{v_F}^2}{G_F}$ where v_f is the Fourier transform of the velocity field and G_F is a low-pass filter.

The performances of the combined SIFT-CPD method, for the reference Pulse Number: 82315 are illustrated in Figure 4. The first frame was used as a reference for the rotation correction. The evolution of the variance (provided by the CPD registration method) shows a clear increase of the uncertainty in determining the rotation angle in the last quarter of the video frames. This is related to the variations in illumination of the video scene (probably due to ELMs). As a consequence, different parts of the vessel structure are revealed with different intensities inducing variations in the localization of SIFT key-points. The points for which the variance is above a certain threshold ($\sigma^2 > 0.01$) have been removed and a local regression has been performed (weighted linear least squares regression and a second degree polynomial model). Lower weights are assigned to outliers in the regression. Zero weight is assigned to data outside six mean absolute deviations. 95.9% of the data was used for determining the rotation angle. Its evolution can be modelled using polynomials. Figure 5 shows the effect of the rotation correction.

The selection of the reference frame is an important input parameter. An illustrative example is presented in Figure 6 for the video Pulse Number: 81980. As the images may change significantly for the frames composing a video, different reference frames must be used for different regions of the video. An indicator for identifying significant changes in the appearance of images is the Tsallis entropy (TE) [15] which has been successfully used for the detection of image vibrations in Ref. 3. However for videos characterized by very low luminosity, TE variations fall below the statistical uncertainty. The automatic identification of the moment when a new reference frame is needed is still under investigation.

The method proved to work well for a high level of noise in the images. As the noise determines the occurrence of false matching SIFT descriptors, the variance of the retrieved rotation increases depending on the amount of noise. For videos characterized by a very low luminosity (5% from the whole database), the rotation evolution has been retrieved only after applying a contrast adjustment. This implies the introduction of an adjustable parameter, which, in principle, has different values for different videos. Parasitical SIFT descriptors may occur together with the contrast enhancement, leading again to an increased variance characterizing the rotation angle.

3. CONCLUSIONS

The image registration methods, based on SIFT descriptors and CPD registration, for rotation correction, provide good results. In most cases the method uses the original input images, without any pre-processing step. The parameters of the algorithm are the same for all videos in the database. The algorithm incorporates several procedures for rejecting outliers. In most cases the evolution of the rotation angle can be approximated by a polynomial fit. The method is able to deal with extremely noisy images and it provides good results for images characterized by very low luminosity. The computation time is compatible with the inter-shot analysis.

ACKNOWLEDGMENTS

This work, supported by the European Communities under the contract of Association between EURATOM and MEdC, ENEA and CCFE, was carried out under the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

[1]. L.G. Brown, ACM Computing Surveys **24**, 326 (1992).

[2]. B. Zitova, et al., Image and Vision Computing, **21**, 977 (2003).

[3]. A. Murari, et al., IEEE Transactions on Plasma Science, **41(10)**, 3030 (2013).

[4]. R.N. Bracewell, The Fourier Transform and Its Applications, McGraw-Hill, New York, 1965.

[5]. V. Martin, et al., Review of Scientific Instruments **81**, 10E113 (2010).

[6]. E. Gauthier, et al., Fusion Engineering and Design, **82(5–14)**, 1335 (2007).

[7]. G. Arnoux, et al., Review of Scientific Instruments **83**, 10D727 (2012).

[8]. J.A. Alonso, et al., 34th EPS Conference on Plasma Physics, Warsaw, 2–6 July 2007 ECA, Vol. 31F, 2007, p. 2.124.

[9]. E. de la Cal, et al., 39th EPS Conference on Plasma Physics 16th International Congress on Plasma Physics, 2 – 6 July 2012, Stockholm, Sweden.

[10]. D.G. Lowe, International Journal of Computer Vision, **60–2**, 91 (2004).

[11]. T. Lindeberg, Journal of Applied Statistics, **21–2**, 224 (1994).

[12]. T. Tuytelaars, et al., Foundations and Trends in Computer Graphics and Vision, **3(3)**, 177 (2007).

[13]. N. Mori, et al., Journal of Engineering Mechanics, **133(1)**, 122 (2007).

[14]. A. Myronenko, et al., IEEE Transactions on Pattern Analysis and Machine Intelligence, **32(12)**, 2262 (2010).

[15]. C. Tsallis "Possible generalization of Boltzmann–Gibbs statistics" Journal of Statistical Physics, **52**, 479 (1988).

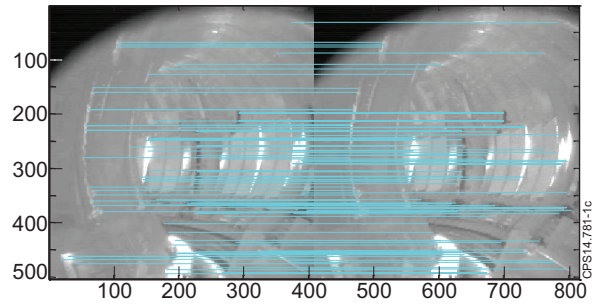

Figure 1: Example of image matching using SIFT descriptors in case of the KL-7 camera (images corresponding to JET pulse 68815). Approximately 180 ÷ 220 key-points descriptors are generated and 90 ÷ 120 matches are found after false match filtering.

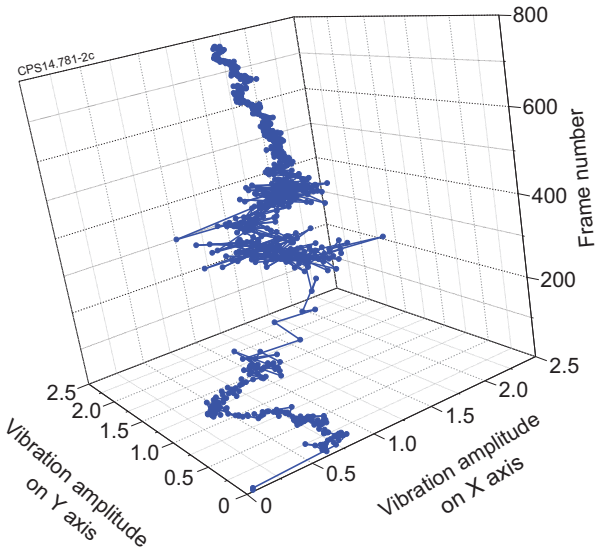

Figure 2: Illustration of the image vibration evolution for a set of 800 frames recorded for the JET Pulse Number 68815.

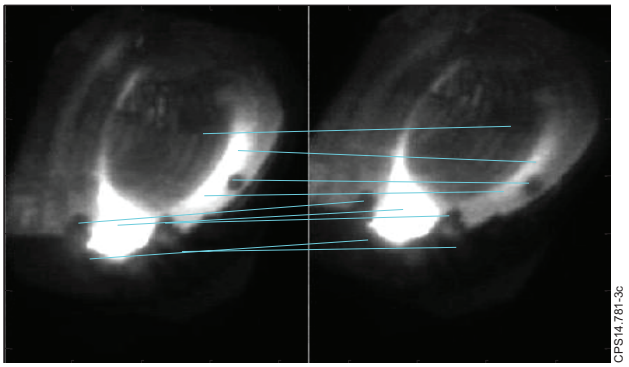

Figure 3: Examples of KL-8 image matching (JET Pulse Number 82315). Equivocal information is provided in a significant number of cases.

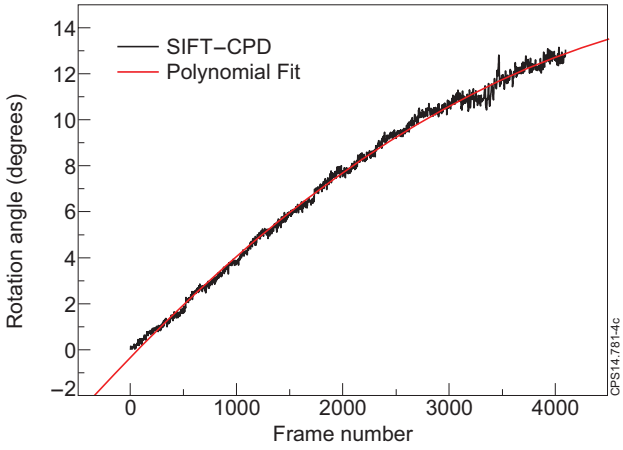

Figure 4: The evolution of the rotation angle (black curve) and its approximation using a 4th order polynomial fit.

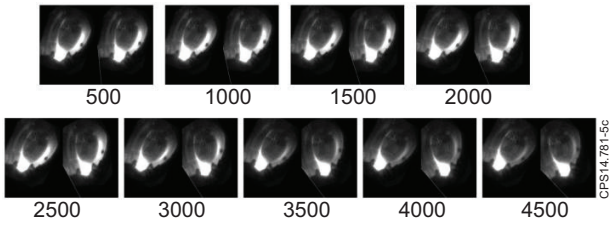

Figure 5: The effect of the SIFT-CPD image correction method for the Pulse Number: 82315 video. For each couple of frames, the one on the left is the registered version.

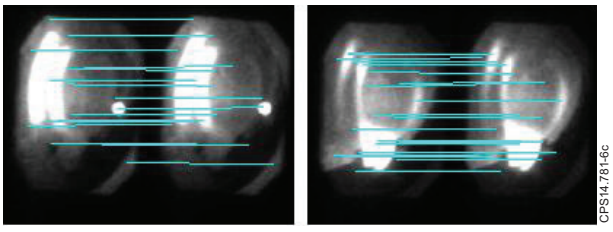

Figure 6: The appearance of the frames may change significantly in time together with the SIFT key-points localisation as illustrated here for the video Pulse Number: 81980.