Nuclear Engineering and Design 335 (2018) 161–171

Contents lists available at ScienceDirect

Nuclear Engineering and Design

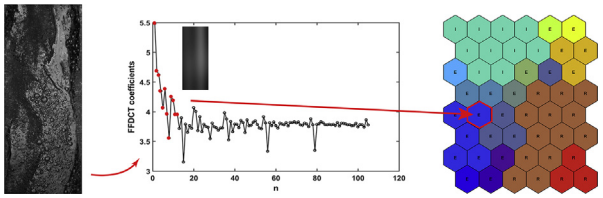
journal homepage: www.elsevier.com/locate/nucengdes

Classification of natural circulation two-phase flow image patterns based on self-organizing maps of full frame DCT coefficients

Roberto N. de Mesquita*, Leonardo F. Castro, Walmir M. Torres, Marcelo da S. Rocha, Pedro E. Umbehaun, Delvonei A. Andrade, Gaiane Sabundjian, Paulo H.F. Masotti

Instituto de Pesquisas Energéticas e Nucleares (IPEN/CNEN-SP), Av. Professor Lineu Prestes 2242, 05508-000 São Paulo, SP, Brazil

GRAPHICAL ABSTRACT



ABSTRACT

Many of the recent nuclear power plant projects use natural circulation as heat removal mechanism. The accuracy of heat transfer parameters estimation has been improved through models that require precise prediction of two-phase flow pattern transitions. Image patterns of natural circulation instabilities were used to construct an automated classification system based on Self-Organizing Maps (SOMs). The system is used to investigate the more appropriate image features to obtain classification success. An efficient automated classification system based on image features can enable better and faster experimental procedures on two-phase flow phenomena studies. A comparison with a previous fuzzy inference study was foreseen to obtain classification power improvements.

In the present work, frequency domain image features were used to characterize three different natural circulation two-phase flow instability stages to serve as input to a SOM clustering algorithm. Full-Frame Discrete Cosine Transform (FFDCT) coefficients were obtained for 32 image samples for each instability stage and were organized as input database for SOM training. A systematic training/test methodology was used to verify the classification method. Image database was obtained from two-phase flow experiments performed on the Natural Circulation Facility (NCF) at *Instituto de Pesquisas Energéticas e Nucleares* (IPEN/CNEN), Brazil.

A mean right classification rate of 88.75% was obtained for SOMs trained with 50% of database. A mean right classificationrate of 93.98% was obtained for SOMs trained with 75% of data. These mean rates were obtained through 1000 different randomly sampled training data. FFDCT proved to be a very efficient and compact image feature to improve image-based classification systems. Fuzzy inference showed to be more flexible and able to adapt to simpler statistical features from only one image profile. FFDCT features resulted in more precise results when applied to a SOM neural network, though had to be applied to the full original grayscale matrix for all flow images to be classified.

* Corresponding author.
E-mail address: rnavarro@ipen.br (R.N. de Mesquita).

https://doi.org/10.1016/j.nucengdes.2018.05.019
Received 12 January 2018; Received in revised form 12 May 2018; Accepted 16 May 2018
Available online 25 May 2018

1. Introduction

Natural circulation phenomenon has been used for many years as a cooling mechanism in safety systems of nuclear power plants. These cooling systems are being included in new reactor design projects as main or redundant heat removal mechanism. It is very important to improve knowledge about heat transfer behavior of two-phase flow regimes in these new reactor designs. Natural circulation instabilities characterize important limiting conditions of these cooling systems, where heat transfer parameters such as critical heat flux are studied (Nayak and Sinha, 2007).

Accident analysis in nuclear reactors were improved when computational resources were available, and research was mostly focused on thermal-hydraulic instabilities. These studies were initiated by Ledinegg in 1938 (Ledinegg, 1938) and developed during the 60's when the high-power-density boilers and boiling water reactors (BWR) were developed (Ruspini et al., 2014). Two-phase flow instabilities study (Chiang et al., 1994) is mainly related to nuclear safety. In cited recent review, Ruspini et al. (2014) describe a "new' kind of flow instability that is called Natural Boiling Oscillations (NBO) by Chiang et al. (1994). NBO is caused by the accumulation of vapor downstream from the heated section with a periodic expulsion of this vapor causing water to fill the adiabatic section and causing a periodic phase shift of about 180° between the flow rate and the pressure drop oscillation in the region where the vapor accumulates. Other two-phase flow instability described by Ruspini is called Geysering, described in upward vertical boilers and occurring for low power and low flow rates. It is referred as having three stages: boiling delay, condensation (or expulsion of vapor) and liquid returning.

Two-phase flow studies using image processing and intelligent algorithms have been recently published (de Mesquita et al., 2012; Fu and Liu, 2016; Pouryoussefi and Zhang, 2015) in search for ideal techniques to promptly identify flow patterns and associated parameters which should enable better flow estimates and control. de Mesquita et al. (2012) have classified different natural circulation instability stages through a Fuzzy Inference System based on statistical properties of detected peaks on different grayscale image profiles. Satisfactory results were obtained for pattern detection based on small quantity of input data extracted from high-resolution grayscale images of two-phase flow natural circulation experiment.

Seleghim and Hervieu (1998) proposed a relation between flow type transitions and time-frequency covariances of void fraction signals. Neural networks have been used to detect phase transitions based on signal changes by the same group more recently (Barbosa et al., 2010).

Many different techniques have associated qualitative image analysis with two-phase flow parameters study, including flow-type transitions, void fraction, dry angles and others (Hsieh et al., 1997; Kattan et al., 1998; Kirouac et al., 1999; Maurus et al., 2002; Shamoun et al., 1999). Pouryoussefi and Zhang (2015) developed a classification system for classical two-phase flow patterns based on Fuzzy Logic an Genetic Algorithm applied to images generated by four different turbulent models. Watershed segmentation associated with an adaptive threshold enabled bubble number density estimation for air-water two-phase vertical upward flow (Fu and Liu, 2016). Other recent related study identifies two-phase flow stages of solid-liquid experiments on Particle Image Velocimetry (PIV) images using image processing techniques (Shi and Zhang, 2015).

In this paper, flow patterns associated with natural circulation instabilities are classified using a neural-network-based algorithm. These instability stages were denominated based on classical Bouré classification (Boure et al., 1973). Two-phase flow patterns associated with natural circulation instabilities were more recently reviewed (Delhaye et al., 1981; Nayak et al., 2007; Ruspini et al., 2014). "Chugging" is the usual term to denominate the characteristic periodic expulsion of coolant from a flow channel.

Periodic two-phase flow instabilities have been studied through the Natural Circulation Facility (*Circuito de Circulação Natural* – CCN) installed at *Instituto de Pesquisas Energéticas e Nucleares*, IPEN/CNEN (Andrade et al., 2000; de Mesquita et al., 2012; de Mesquita and Penha, 2010). This facility is an experimental circuit designed to provide thermal-hydraulic data related to one and two-phase flow under natural circulation conditions. One of the primary features of this experimental circuit is to allow extensive visualization of flow through its tubes which are all glass-made.

This work uses the same image database used in previous work to enable comparison of classification efficiency between a previously developed Fuzzy Inference System (de Mesquita et al., 2012) and Self-Organizing Maps methods. The Fuzzy classification mainly used statistical properties of grayscale profiles that were extracted from correspondent images to each instability stage.

During the development of the present classification system using SOMs, it became evident that previously used statistical grayscale profile features were not enough for the SOM classifier to make a clear distinction among the instability patterns. Simple variance measures including the whole image were not able to characterize each stage. The heterogeneous shapes of Expulsion stage were many times confused with the other two stages studied. The paper was initially based on the same features used on Fuzzy System, which were applied in that work over single grayscale profiles of two-phase flow studied images. The features used in that paper, although statistic in nature, were related to a specific grayscale profile. Fuzzy methodology showed to be more appropriate to treat and adjust fuzzy rules to separate each image class. However, the Fuzzy classifier was high dependent of the chosen profile. SOM was not able to use these features as good classifier inputs. New features were necessary to be tested to improve classification success. Full-Frame Discrete Cosine Transform (FFDCT) coefficients showed to be adequate components of a feature vector which was very efficient in characterizing the images corresponding to each instability stage studied. The next sections will explain the methods used to develop this classification system and how the test results were statistically studied to estimate the classifier efficiency.

1.1. Full-Frame Discrete Cosine Transform (FFDCT)

Discrete Cosine Transform (DCT) is a frequency domain transform which is part of Discrete Fourier Transform (DFT) that has been used since the 70s by Ahmed et al. (1974) in applications such as digital processing and pattern recognition and Wiener filters (Pratt, 1972). During the 80s, Chen and Pratt (1984) used DCT to digital image compression obtaining good results.

DCT usual application is done through block-based transform coding to obtain lossy data compression. A compression scheme based on DCT was standardized by the Joint Photographic Experts Group (JPEG) of the International Standardization Organization (ISO). This technique obtains high compression ratios at the expense of producing some artifacts induced by block decomposition at compression ratios of 8–12:1 (Beretta et al., 1994). The JPEG scheme uses DCT through image blocks of 8 by 8 pixels. These blocks are used for image reconstruction after they are transmitted and received. Artifacts could become visible with this technique, and for this reason, this compression is undesirable or unacceptable for medical applications (Villasenor, 1993). A more recent and reliable DCT technique consists of applying DCT to the entire image as a single block. This technique is usually cited as Full-Frame DCT (FFDCT). Although it is more computationally intensive to perform, it yields better image quality results at higher compression ratios (Cavalaris, 1998).

In this paper, FFDCT was used to obtain compact information that should represent the main morphological characteristics of each flow pattern studied. FFDCT is mainly used in medical images (Wilhelm et al., 1991) and to produce watermarks in digital images (Emami, 2016). To enable watermarking applications, FFDCT coefficients probability density functions can be modeled without loss of

performance by a Laplacian density function with variance decrease with frequency (Barni et al., 1998).

FFDCT coefficients showed to be an important feature to be used on classification tasks which depend upon image representation. The FFDCT ability to compress most of image frequency information in few ordered coefficients that represent the main image characteristic of the studied instability stages. The FFDCT coefficients ordering is such that low frequency information and high frequency information extracted from image can be accessed depending on which frequency band of information is chosen to be used. In this work, the first ordered coefficients used were the ones correspondent to low frequency information extracted from analyzed images. This aspect will be better described in following sections.

2. Methodology

The main purpose of the classification task was to automatize the recognition of images pertaining to one of the three stages of the 'chugging' cycle. These stages are cyclical with regular periods and are called as Incubation (I), Expulsion (E) and Refill (R) as described by Boure et al. (1973).

The algorithm was used to associate each instability stage of natural circulation cycle imaged at a Natural Circulation Facility (NCF) to one or a set of image features. This association was implemented by a Self-Organizing Map neural network trained based on a previous database (de Mesquita et al., 2012).

Two-phase flow images were acquired concomitantly with temperature measurements to stablish the time-intervals in which each instability stage was being imaged. Based on these time-intervals, image database for each instability stage was composed of heterogeneous image patterns.

2.1. Natural circulation facility

2.1.1. Experimental circuit

The experimental images used in this work were acquired at Nuclear Engineering Center (CEN) of the Nuclear Energy Research Institute of São Paulo, IPEN. The images were captured during two-phase flow experiments performed at the Natural Circulation Facility (NCF) (Fig. 1) where two-phase flow instabilities were generated. These natural circulation instabilities are experimentally verified to behave cyclically at determined time periods. This cycle can be divided into three stages. Each of these stages have their own period, with defined transitions between them. This experiment is better described in previous work (de Mesquita et al., 2012).

The NCF was designed to provide thermal hydraulic data related to one and two-phase flow at ambient pressure. It was all constructed of glass tubes enabling flow patterns visualization through all circuit. The circuit is detailed described elsewhere (Andrade et al., 2000; de Mesquita et al., 2012). Image acquisition (Fig. 1(b)) was done above electric heater region (Fig. 1(a)).

2.1.2. NCF instability cyclic behavior

The NCF presents two-phase periodic instability with well-defined period of oscillation. This instability has been described as 'chugging' (Andrade et al., 2000; Boure et al., 1973; de Mesquita et al., 2012). This instability is characterized by periodic expulsion of coolant from the channel. The two-phase flow experiments at NCF are usually adjusted to sustain a cyclic and periodic behavior of this instability.

The overall chugging cycle, observed at NCF experiments and used for this work, was obtained with a 7270W heating power in a 9500s test. Images were acquired with an interval of 1s during all experiment. Periodic transition of stages was evaluated by measuring time intervals between images corresponding to stage transitions. The heating power was estimated to be raised up leading the flow temperature up to 98 °C with an ambient temperature of 25 °C. Cooling flow rate of 140 L/h was kept constant during the experiment.

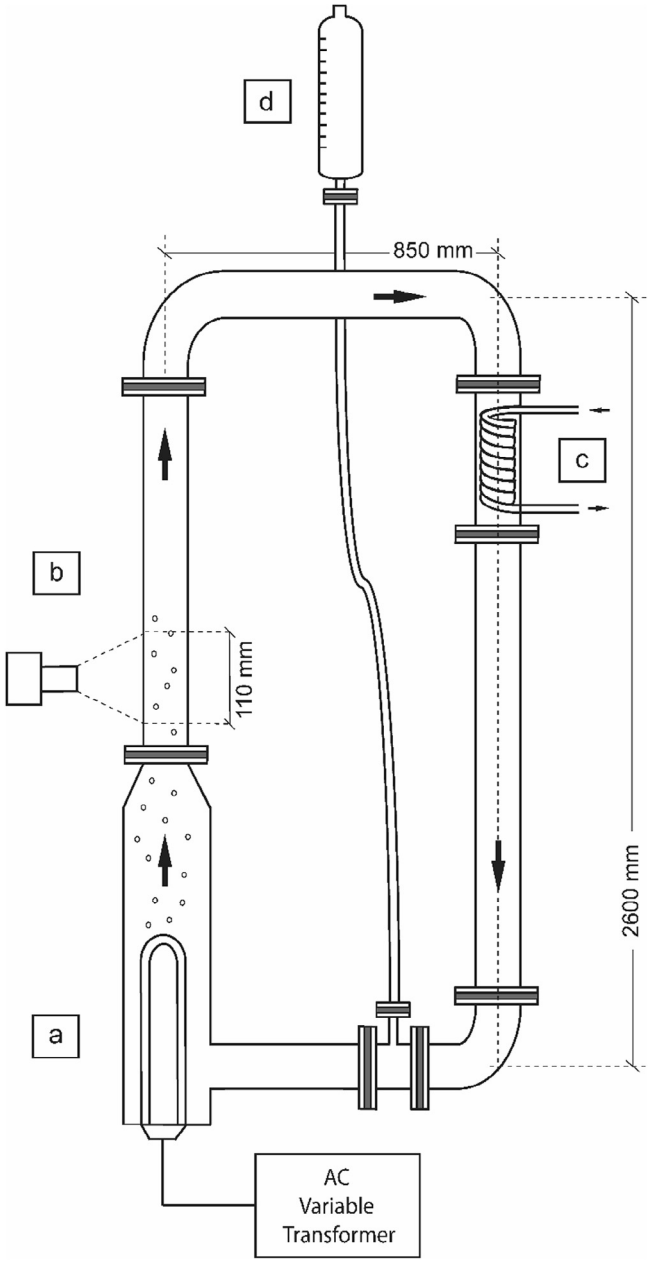


Fig. 1. Schematics of NCF: electrical heating section (a), visualization section (b), cooling section (c) and expansion tank (d).

The image database was composed of selected images corresponding to each instability stage. The selected images for each stage were chosen to be in the middle of time interval between stage transitions (Fig. 2). Fig. 3 presents examples of image database samples obtained for each chugging cycle stage.

The incubation stage is characterized by the growth of number and size of vapor bubbles and accumulation of vapor at upper horizontal region of the circuit. Pressure grows and flow rate decreases causing expansion tank to be filled with liquid from the cold leg. Expulsion stage starts when slug flow begins to be replaced by churn flow when liquid entrained by vapor is expulsed from hot leg. Expansion tank attains its maximum value at this stage. Finally, the Refill stage begins when the flow rate direction is inverted by the difference of hydrostatic head. Here, the hot water at the heater is replaced by cold water coming from cooler. The vapor production at the heater decreases and the horizontal part of the hot leg is filled with water again, beginning the

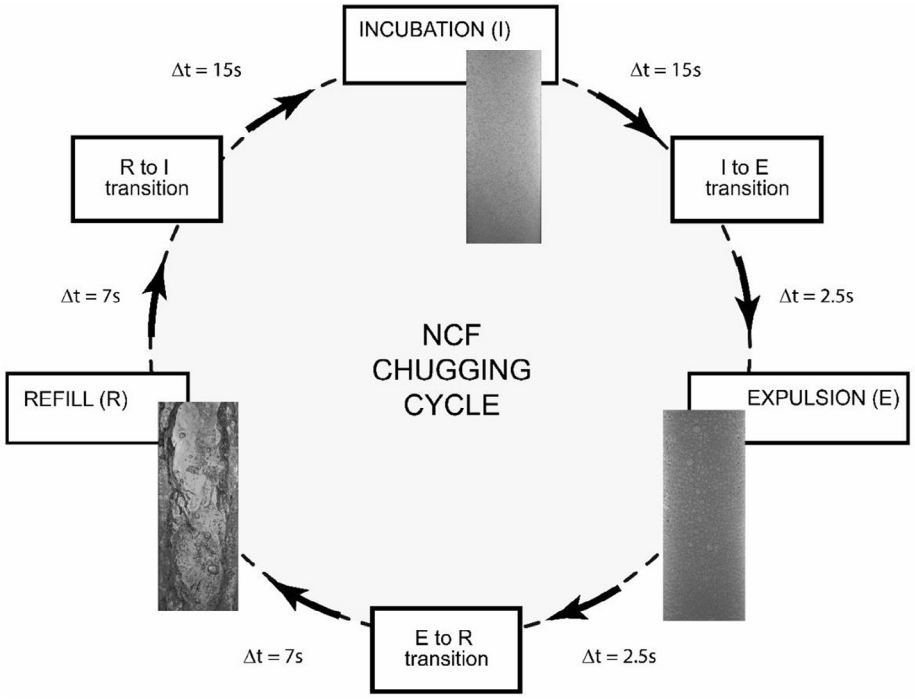


Fig. 2. Graphical representation of NCF instability cyclic behavior.

overall cycle. The cycle repeats itself when hot leg is once more filled with water. The regular T period for this experiment was evaluated based on temperature minimum-peaks time interval. Image patterns presented the same cycle period. This periodic flow oscillation behavior can be thoroughly observed in this facility due to its glass-made tube structure.

A graphical representation of chugging cycle stages at NCF is shown in Fig. 2 where the three main stages are represented. Total cycle time is 49 s and each stage is represented in this figure. Each block represents a typical vertical section image. The blocks are connected by arrows indicating the sequence order of stages and the interval time between the moment when images were taken. The regular T period of 49 s for a complete chugging cycle is estimated after two-phase flow stabilization occurs. The cycle is composed of an Incubation stage (T to (T + 30)s), an Expulsion stage ((T + 30)s to (T + 35)s) and a Refill stage which lasts for the remaining 14 s of the cycle period.

2.2. Image acquisition

Image acquisition was done by using high-resolution digital camera with 250 μs shutter speed. The captured tube section was 20 cm long of a 3 m vertical glass column. Lens mount was configured to enable macro focus and image acquisition was done at approximately one frame per second rate during different cycles of 1000–1500 s long. Typical acquisition modes generated 3888 × 2592 pixels of a longitudinal tube section with a resolution of approximately 0.03 mm/pixel. Backlight illumination technique showed to be the optimal condition to obtain image borders best definition. Images were acquired at an approximate 200 mm longitudinal section of the cylindrical hot leg tube (46.3 mm external diameter) shown on Fig. 1(b).

Image Database was organized based on the three main chugging stages. The database was composed of selected images related to each stage of chugging cycle. From 2530 images, 32 sample images were selected to characterize each flow stage. The images in Fig. 3 show four examples for each chugging stage. From these images, it is possible to observe that there are visual similarities and differences among the same stage examples.

Pattern images were acquired simultaneously with circuit temperature measurements synchronized in time. Periodic behavior was confirmed by the detection of the refill-to-incubation stage transition image pattern. This detection is described with more detail in previous works (de Mesquita et al., 2012; de Mesquita and Penha, 2010).

The training and test image files were organized in three different folders corresponding to each instability stage class. Each class folder contained 32 full-sized 10^7 pixels RGB images (red-greenblue pattern) in compressed file format. Images were pre-processed with an interpolation done by the gray-level transform Matlab function '*rgb2gray*'.

The images were stored using the better compression JPEG quality available on the high-resolution Canon® EOS-5d camera. This compression quality file format has very small information losses in the high frequency domain and in color information. As our method is based on a low frequency feature and is applied over grayscale images, this loss would have negligible influence. There are advantages in speed of image acquisition and image processing.

2.3. Full-Frame Discrete Cosine Transform (FFDCT) feature extraction

The image features were mainly extracted based on Full-Frame Discrete Cosine Transform (FFDCT) to classify each instability stage. This feature was chosen after the unsuccessful use of statistical properties as input to SOM neural network classifier. Statistical properties associated to line profiles were previously used to classify instability stages (de Mesquita et al., 2012).

Images were chosen at the middle of each chugging stage cyclic time. Therefore, for the same stage, there were different image samples with variations of image patterns. This work searched for more perennial image features that could be present in the same instability stage image database (Fig. 3).

FFDCT showed to be the more appropriate and efficient image feature to represent the common pattern for the same instability-stage images. Special importance was necessary to be given to Expulsion stage, which presented images with different patterns but almost always presenting big volumes of vapor inside the imaged section. These big void sections inside images were easily detected by FFDCT coefficients, as can be seen in Fig. 4. This figure shows that with few FFDCT

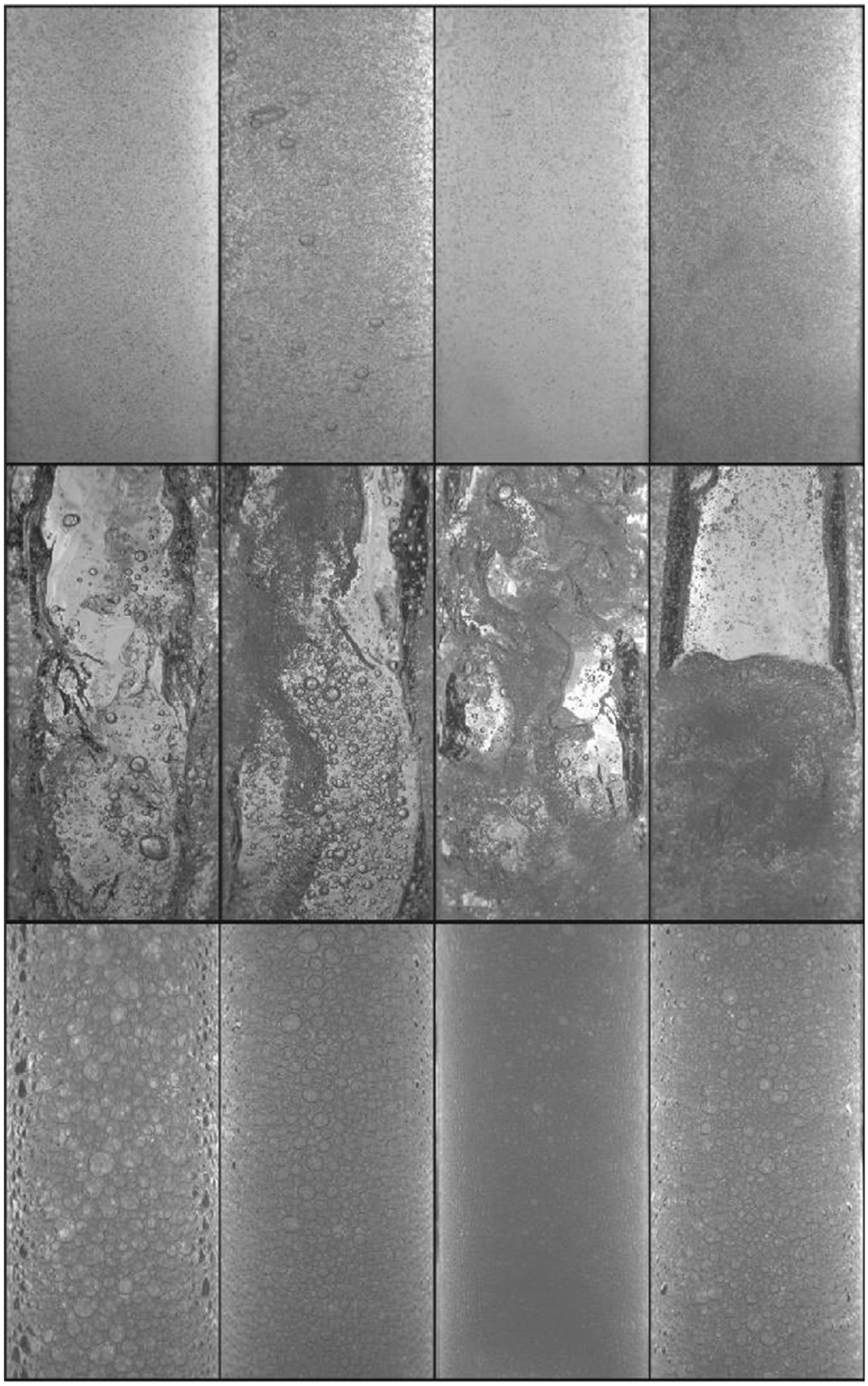


Fig. 3. Cropped images of NCF chugging cycle: Incubation (I) samples at the first row, Expulsion (E) samples at the second row, and Refill (R) samples at the third row.

coefficients it is possible to reconstruct the main void areas with big regions of white pixels. A typical FFDCT vector with 100 coefficients obtained for NCF analyzed images is presented in Fig. 5.

The low frequency information present in the image (regions with smoother gray level transitions) is registered in the first coefficients. These great areas presenting clearer gray tones could be preserved using only 12 initial coefficients using the *zig-zag* technique (Eq. (4)). These transforms have a strong power of information compaction.

$$\mathrm{M}=\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \rightarrow \text{ coef} = [\mathrm{a\ b\ d\ g\ e\ c\ f\ h\ i}] \tag{4}$$

One the main objectives to use FFDCT was to characterize the global blur associated with the big regions of vapor in two-phase flow instability stage called Expulsion as can be seen in Fig. 4.

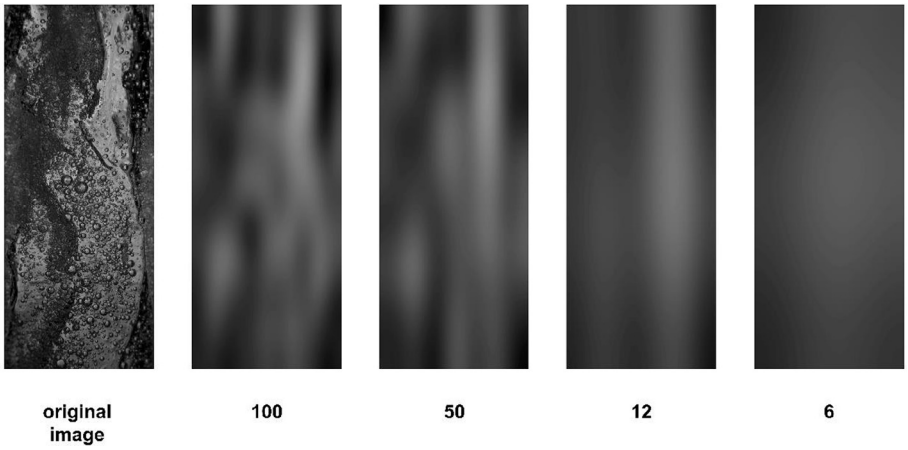


Fig. 4. Typical Expulsion stage original image, with respective reconstructed images using 100, 50, 12 and 6 full-DCT coefficients.

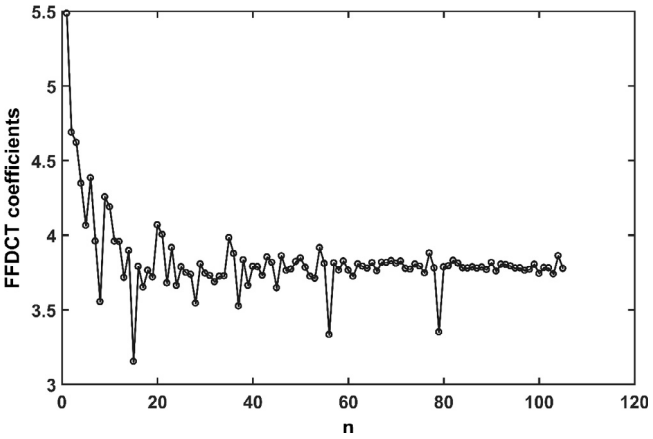


Fig. 5. FFDCT feature vector with 100 coefficients obtained through zig-zag technique.

2.4. Overall classification system

The methodology used to apply SOM to flow images is represented on Fig. 6. Images were acquired as is described in section 2.2 and organized in proper folders to enable appropriate neural network training. A sequence of steps precedes the proper SOM training to classify flow instability stages.

Database was organized aiming a comparison with previously classification methodology using Fuzzy Logic (de Mesquita et al., 2012). Image database was used to create FFDCT vectors for each chugging stage class.

SOM training was done in two different sampling percentage of whole database, 50% and 75% of available images for each instability stage, or 16 and 24 images from the 32 images for each instability stage. In each sampling percentage for training, the rest of the data were used to test the classification task (Fig. 6) where TR% is the portion of data samples used to train the SOM neural network. The *DCT2* function from Matlab (Mathworks Inc, 2015) was applied to each image sampled for training and a FFDCT coefficient matrix was obtained. This function applies two-dimensional discrete cosine transform to image matrixes. Through *zig-zag* technique, the first 12 coefficients (corresponding to low frequency) were chosen to form a characteristic vector for each sampled image. An example of a FFDCT vector with 100 coefficients is show in Fig. 5.

The implemented classification task was based on different sampling experiments, enabling a procedure similar to cross-validation technique where random proportional sampling with reposition was done to train the SOM with different sampling.

Final classification tests were performed in different ‘experiments’ with a thousand of repetitions for each sampling type. At first, a thousand tests were done with random initialization for each sampling proportion (50% and 75% from whole database). After each of the 1000 times in which the training was done, a SOM map was created and 15 clusters were searched using *k-means* method. After clustering was performed upon the trained SOM, a classification map was constructed to distinguish each instability stage class. Initially the Best Match Units (BMUs) for each sample image were obtained and an initial labelled map was created (Fig. 8a). Thereafter, a classification SOM was fully labelled (Fig. 8b) with the instability classes chosen for each prototype generated by SOM training process. Based on this last map, a test was performed searching for the nearer (based on Euclidean distance) generated SOM prototype, and its associated label was used to classify the test sample as I, E or R. All test vectors were tested for each of the 1000 random sampling ‘experiments’. The right classification rate was estimated to each flow instability stage and a global classification rate was evaluated.

Representative histograms were generated to show ‘experiments’ right classification distribution. Figs. 11 and 12 show the obtained histograms for each instability stage class and a global Right Classification Rate.

2.5. Self-organizing maps (SOMs)

SOM neural network is a clustering algorithm developed by Teuvo Kohonen (Kohonen, 2001, 1982), based on self-organization of multi-dimensional prototype vectors which are molded and approximated using Euclidean distance among vectors. A two-dimensional map containing these trained prototypes is obtained. The two-dimensional self-organization and training is based on a competitive learning paradigm, where each prototype vector compete to better represent input data ($\overline{x}_j$) distribution. Each prototype $\overline{p}_j$ is represented in a two-dimensional cell which is a unit of a grid of m cells (or neurons), where $\overline{x}_j$ and $\overline{p}_j$ are n-dimensional vectors:

$$\overline{x}_j = \{z_1, z_2, \ldots, z_n\}, z_i \in \mathscr{R}, \tag{1}$$

$$\overline{p}_j = \{t_1, t_2, \ldots, t_n\}, t_i \in \mathscr{R}, \tag{2}$$

Prototype-vectors are competitively trained based on Euclidean distance (Eq. (3)):

$$d(\overline{x}_j, \overline{p}_j) = \sqrt{(z_1 - t_1)^2 + (z_2 - t_2)^2 + \ldots + (z_n - t_n)^2}, \quad j \in [1,2,\ldots,m] \tag{3}$$

The neuron containing the nearest prototype is usually called ‘Best Match Unit’ (BMU) and is chosen to be changed to minimize the distance to the input vectors. The neurons set constitutes a bi-dimensional lattice map which juxtaposes similar prototype vectors. This

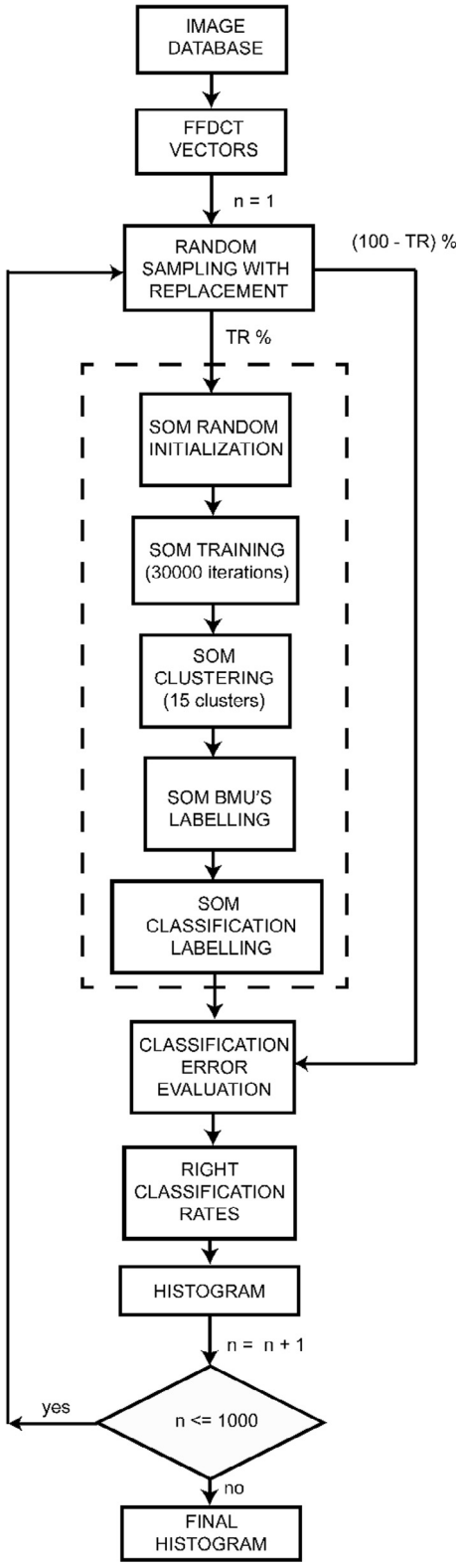


Fig. 6. Instability Flow-type Classification System (IFCS) based on SOM clustering. TR% is the portion of data used for training (50% and 75%).

topographic map preserves implicit statistical information contained in input data. These maps can be considered as a linear generalization of principal component analysis heuristic and when proper training is used, the hexagonal grid of neurons converges (Bouton and Pagès, 1993; de Mesquita et al., 2004).

The map is considered well trained based on quantization error (q_e) which evaluate average distances between each vector and its correspondent BMU. The topographic error (t_e) is another important training parameter which measures the average Euclidean distance between first neighbors' cells on SOM grid. The SOM algorithm was implemented using Som Toolbox 2.0 (LCIS, 2011) over Matlab (Mathworks Inc, 2015) platform.

Table 1
SOM right classification ($\overline{R}$) statistics for 50% sampling.

Classification Pattern	$\overline{R}$ (%)	R_{min} (%)	R_{max} (%)
I	98.74	70.63	100
E	75.61	41.25	98.75
R	91.90	33.75	100
Global	88.75	67.92	99.58

Table 2
SOM right classification ($\overline{R}$) statistics for 75% sampling.

Classification Pattern	$\overline{R}$ (%)	R_{min} (%)	R_{max} (%)
I	98.01	70	100
E	85.82	42.5	100
R	98.12	30	100
Global	93.98	68.33	100

2.5.1. Statistical classification test methodology

The proposed classifier was tested using a statistical procedure using repeated random sampling with reposition. This procedure is one of the cross-validation variants usually applied which requires smaller database size (Kohavi, 1995). All training and classification steps were repeated 1000 times.

The sequence of steps used to test the classifier with statistical sampling described in Fig. 6 is summarized in eleven steps:

i) Database organization based on instability stage in folders containing 32 different images with an acquisition time corresponding to the middle of each chugging instability stage (I, E or R);
ii) FFDCT coefficients evaluation for each database image and storing each of these sets on files inside a correspondent instability stage folder;
iii) Random sampling with replacement of a percentage of FFDCT files to train the SOM and separating the resting samples for testing;
iv) SOM random initialization for each of the 1000 times that a new random sampling is done;
v) SOM training using 30,000 iterations;
vi) Map clustering using 15 clusters as a basis to find groups of similar vectors using the k-means algorithm through *kmeans_clusters.m* function from SOM Toolbox (LCIS, 2011);
vii) Clustered map labelling using BMUs for each training sample;
viii) Map clusters filling with labels chosen based on the class associated with more number of BMUs inside each of the 15 clusters;
ix) Comparison of each test vector with prototypes of the map, and associating the label of winner prototype (Euclidean nearer);
x) Right classification rates obtainment for each class and global result;
xi) After 1000 sampling tries a histogram is constructed.

Based on Kohonen criteria (Kohonen, 2001) and taking into account the number of training vectors (considering 25% of total samples, 24 vectors), a small map should have seven cells (seven prototypes) and a big map should have 113 cells. For a 50% sampling these numbers would double with a map size varying from 14 to 226 cells. For a 75% sampling the map size should stay between 28 and 452 cells. Based on many tries, this work found the ideal size to be 48 hexagonal cells in order to compare the different experiments with different sampling and

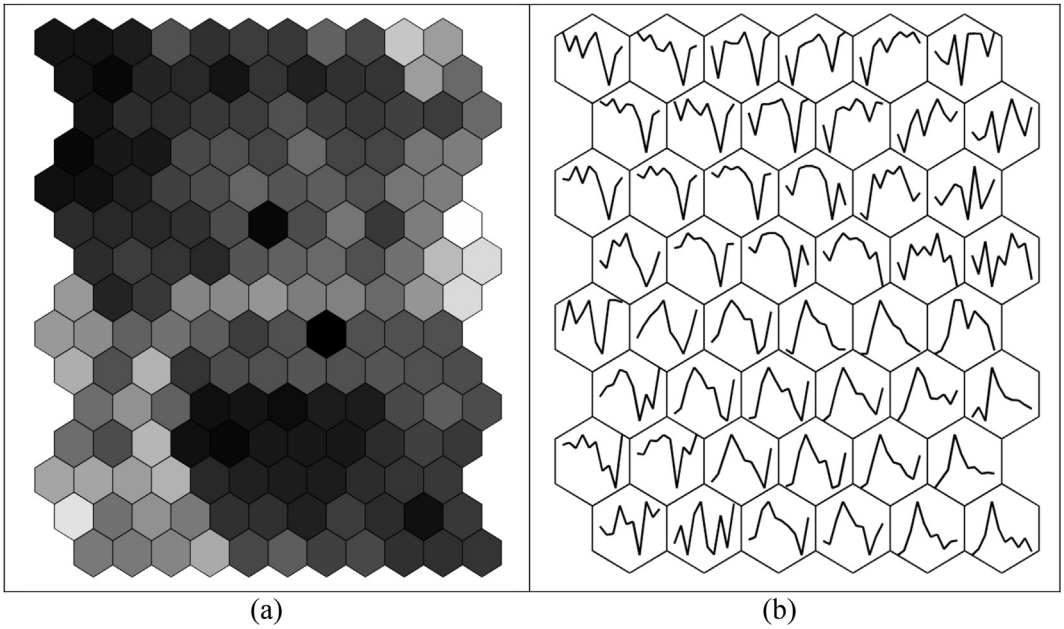


Fig. 7. SOM distance map (a) and FFDCT prototypes (b) generated for one training sample 'experiment' for 50% sampling. Darker cells on distance map (a), represent smaller Euclidean distances between each pair of prototype cells (b).

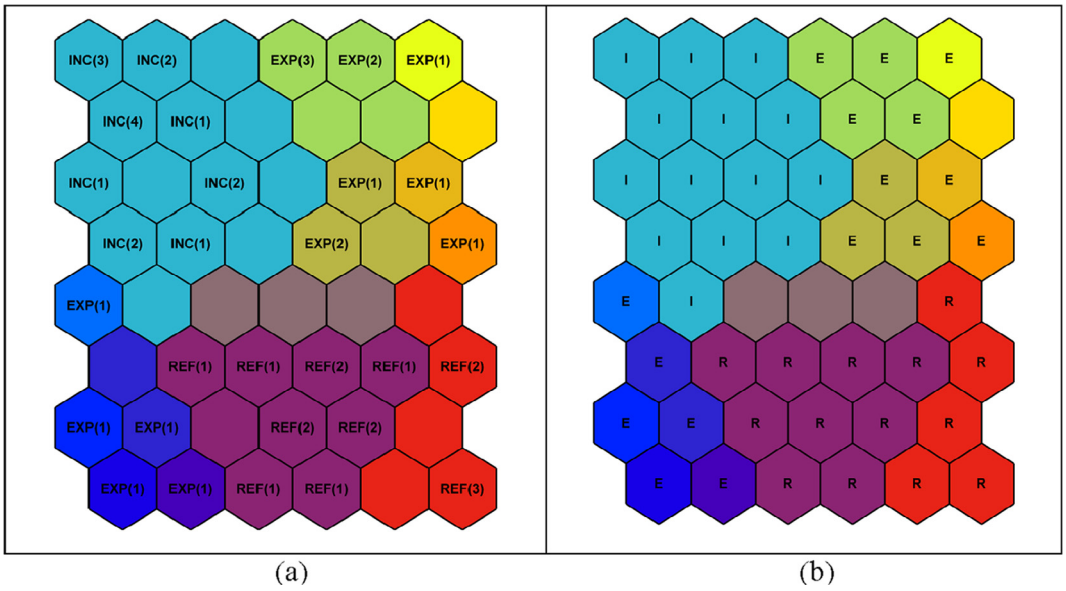


Fig. 8. SOM BMUs labelled map (a) and fully labelled classification SOM map (b) for 50% sampling. Colors represent the 15 clusters used to organize the map.

initialization. The map was trained using 30,000 iterations as basis for all 'experiments'.

3. Results and discussion

The classification task was the main purpose of this work. The use of FFDCT coefficients has shown to be an important feature to characterize different flow regimes. Specially for this work, the high classification rates obtained showed conclusively that this system can be very useful to two-phase flow patterns detection.

Tables 1 and 2 show a global 88.75% and 93.98% right classification rate, based on 1000 tries for different random sampling of 50% and 75% of database respectively (which corresponds to 16 and 24 samples of 32 files for each instability stage class). All tries were done based on random initialization for each sample, assuring statistical distribution demonstrated in histograms shown in Figs. 11 and 12.

The maps were trained with a constant 8x6 size using radius of 4 distance units for initial training phase (30,000 iterations) and radius of 2 units for fine-tuning phase (22,500 iterations). Quantization error (q_e), which is the average distance between each training vector and its corresponding BMU, was kept near 0.0912 during all training stages used in this work. The obtained topographic errors (t_e), which are the proportion of all training vectors for which the first and second BMUs are adjacent cells in SOM map, were kept near 0.0208.

Typical obtained profiles of FFDCT for each instability stage class can be observed in Figs. 7(b) and 9(b) where SOM prototypes. Each of these FFDCT prototype maps have is presented besides the SOM unity matrix (U-matrix), showing the Euclidean distance distribution among prototype cells (Figs. 7(a) and 9(a)). The darker gray colors represent smaller distances between each pair of cells. Clearer gray colors represent bigger distances. These distance maps show a preview of possible clusters. Typical obtained clustered SOMs with the BMUs labelled

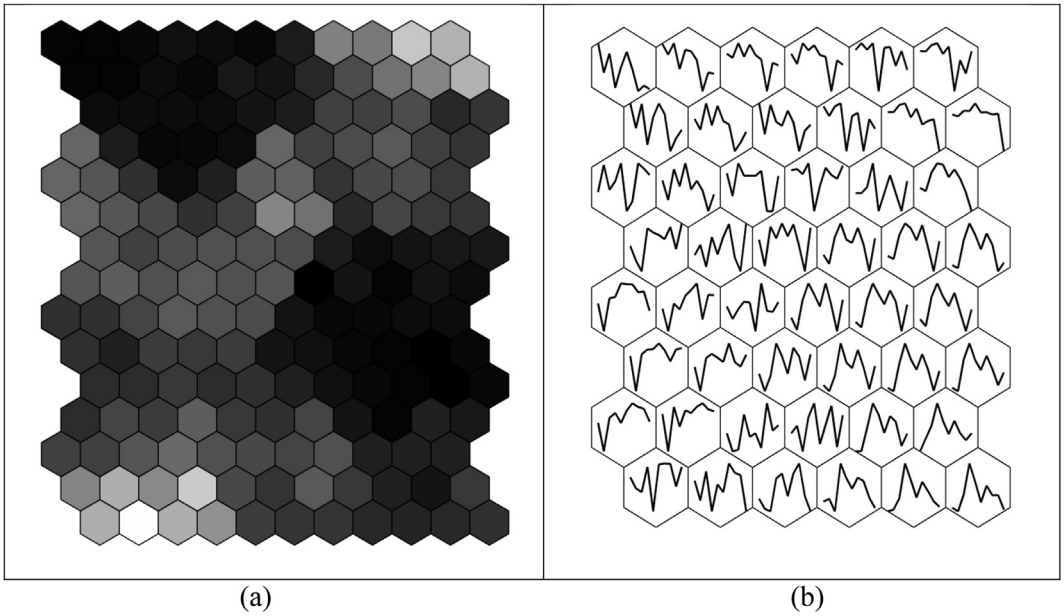


Fig. 9. SOM distance map (a) and FFDCT prototypes (b) generated for one training sample 'experiment' for 75% sampling. Darker cells on distance map (a), represent smaller Euclidean distances between each pair of prototype cells (b).

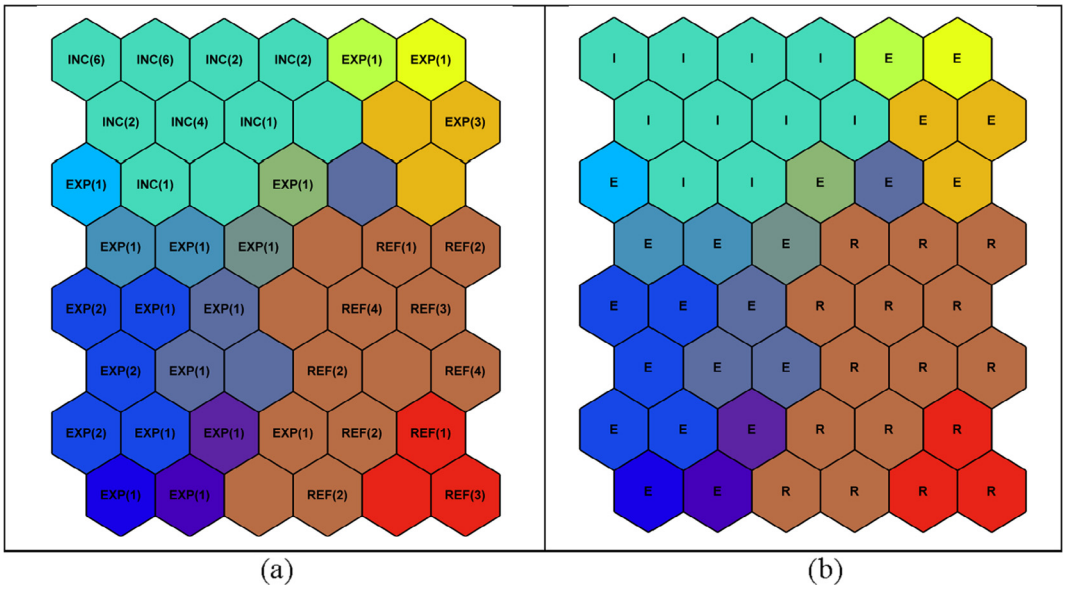


Fig. 10. SOM BMUs labelled map (a) and fully labelled classification SOM map (b) for 75% sampling. Colors represent the 15 clusters used to organize the map.

on it and the corresponding fully labelled SOM used for classification are shown in Figs. 8 and 10 for 50% and 75% sampling proportion. The different colors represented in Figs. 8 and 10 represents the 15 different clusters which were used to classify the map. It is important to emphasize that for each training sample a different map was generated, and all classification procedure was repeated. The maps presented in Figs. 8 and 10 are typical results.

Figs. 7(a) and 9(a) represent the unity matrix (U-matrix) showing the normalized Euclidean distances for all the map cells. Darker grayscale units represent nearer distances between SOM main map cells.

It is possible to observe that the statistical distribution of right classification rates obtained for different sampling proportions (50% and 75%) showed best performance of 75% sampling. The minimum value for both sampling proportion remained within 5 percental points' difference. Expulsion class right classification was the hardest task due to inhomogeneity of image patterns related to this stage, as can be seen in Fig. 3, second line, where four different samples are shown. It can be observed that, although the number of input training samples were always kept constant for each instability sample, the prototypes classified as Expulsion were greater due the inhomogeneity of patterns for the same instability stage.

The classification results are presented in Tables 1 and 2. These tables present the mean right classification rate $\overline{R}$, the minimum right classification R_{min} and the maximum right classification R_{max} for 1000 trials. Table 1 presents the 50% sampling results and Table 2 presents the 75% results.

These results should be considered in comparison with previous Fuzzy System (de Mesquita et al., 2012). The main aspects of these comparisons are: 1) many experiments using simple statistical measures over all image were tested as input to SOM neural network, and were not enough for good classification results; 2) these experiments were

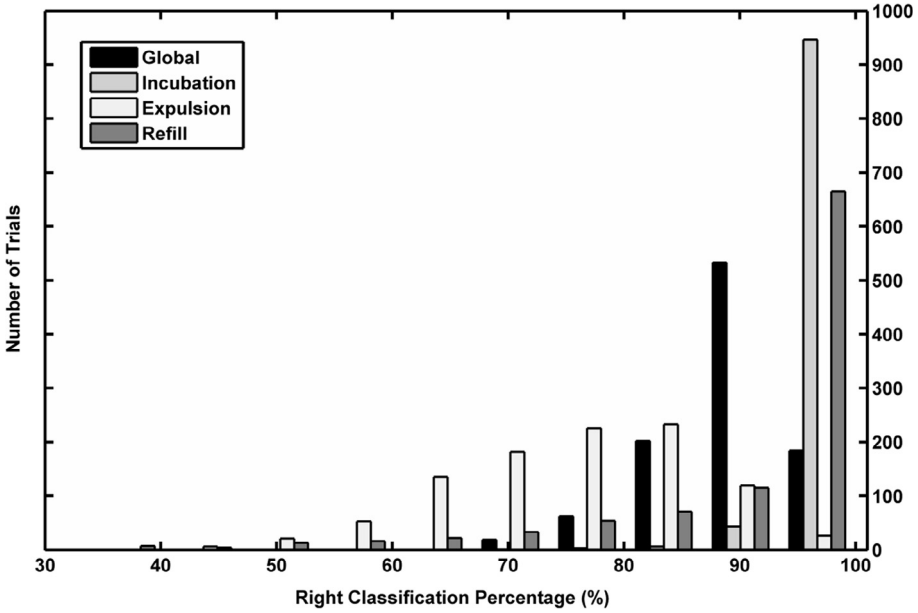


Fig. 11. Right Classification Histogram for 1000 random sampling and initialization for 50% sampling.

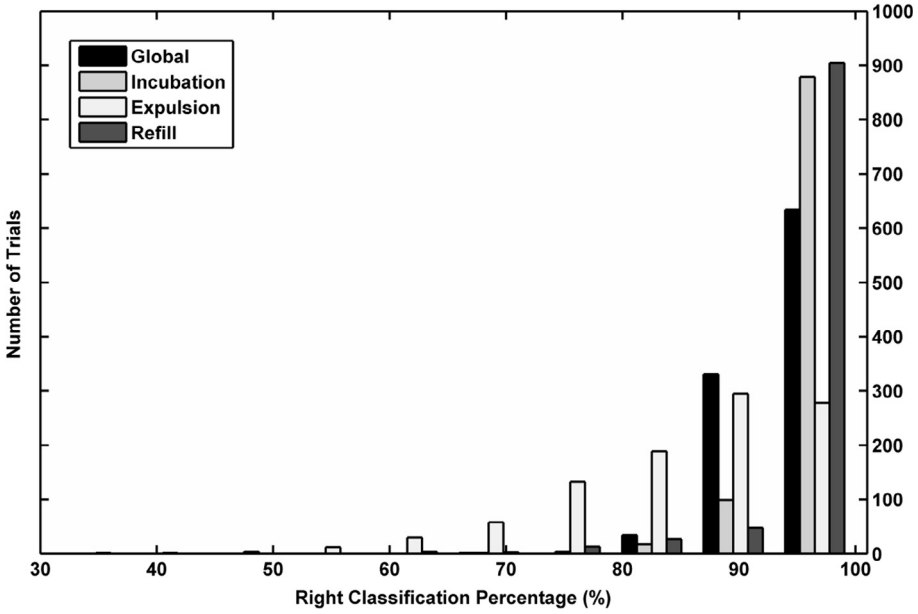


Fig. 12. Right Classification Histogram for 1000 random sampling and initialization for 75% sampling.

not included in the present paper because presented very unsatisfactory results; 3) the Fuzzy classification system used in previous paper was based on two different statistical features applied over a grayscale profile; 4) these profiles were also tried as input to SOM classifier and also presented bad results; 5) Fuzzy classifier had to rely on grayscale profiles, and had to adjust fuzzy rules to two different extracted features simultaneously; 6) in Fuzzy classifier the results depended on which profile was chosen. The FFDCT coefficients showed to be an important image feature that was the unique feature used as input.

4. Conclusions

A new classification technique using the Kohonen neural network (Self-Organizing Maps) was presented using Full-Frame Discrete Cosine Transform (FFDCT) coefficients as input features. These coefficients were extracted from whole images taken experimentally from a two-phase flow experiment performed at a Natural Circulation Facility. The images were taken from natural circulation chugging instability generated in a controlled experiment taking temperature measurements to characterize the oscillation period of the periodic instability.

A methodology for two-phase flow instability classification was previously published (de Mesquita et al., 2012) using a Fuzzy Inference System based on statistical features from single image profiles. The initial purpose of the work was to compare both methodologies. The neural network showed lower capacity to classify these image patterns based only on statistical features extracted from image grayscale profiles. However, the use of FFDCT coefficients proved to be a valuable tool to be used as input to neural network systems on this classification task. This feature showed superior results in extracting associated prototypes to each flow instability stage.

The images were selected based on the corresponding middle time of each chugging stage considered, presenting variable image patterns

for the same stage, specially the Expulsion stage. The classification success of the technique was due to FFDCT efficiency in capturing the common characteristic presented in different image patterns within the same chugging stage. This was obtained with few 12 coefficients which kept information of great void regions from images that originally were taken with resolution of 3888 × 2592 pixels (10 Mp).

The SOM classification task required that the image extracted features were more concisely representative of each instability stage. The Fuzzy rules seem to be more flexible and able to compensate the heterogeneity of image patterns present in the same instability stage. The present study was able to confirm that it was possible to use as few as 12 coefficients from FFDCT feature to classify the instability stages. These coefficients are already known to be effective in compacting information contained in image, as are usually applied in compacting algorithms. The use of such few FFDCT coefficients to characterize an image pattern used as input to a neural network was one important achievement. It can be applied to two-phase flow studies and can have broader applications as well.

The statistical method used for validation of the classification trials was executed with 1000 tests for each sampling proportion used for training. Results showed right classification mean values near 90%. The histograms showed robustness of results.

The Incubation (I) and Refill (R) patterns were better identified with higher right classification rates as the images pertaining to these classes have more homogeneous morphological characteristics. Expulsion (E) patterns have more inhomogeneous morphological features causing more difficulty to be identified.

These results represent an important contribution to this research area. Monitoring multiphase flow in nuclear systems through digitally registered images or direct visualization it is a key area to be developed both on basic research on thermal-hydraulic heat transfer problems and on industrial applications.

Acknowledgments

The authors would like to thank Luis A. Macedo for his special contribution to NCF experiments. The authors also wish to acknowledge experimental assistance from the Department of Chemical Engineering of the Polytechnic School of the University of São Paulo.

References

Ahmed, N., Natarajan, T., Rao, K.R., 1974. Discrete cosine transform. IEEE Trans. Comput. 90–93.

Andrade, D.A., Belchior Jr, A., Sabundjian, G., 2000. Two-phase instabilities in a natural circulation rectangular loop. In: Proc. 8th International Conference on Nuclear Engineering.

Barbosa, P., Crivelaro, K., Seleghim Jr., P., 2010. On the application of self-organizing neural networks in gas-liquid and gas-solid flow regime identification. J. Braz. Soc. Mech. Sci. Eng. XXXII, 15–20.

Barni, M., Bartolini, F., Piva, A., Rigacci, F., 1998. Statistical modelling of full frame DCT coefficients. In: Signal Processing Conference (EUSIPCO 1998), 9th European.

Beretta, P., Prost, R., Amiel, M., 1994. Optimal bit allocation for full frame DCT coding schemes application to cardiac angiography. In: Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE). SPIE – Int Soc Optical Engineering, PO BOX 10, Bellingham, WA 98227-0010, Newport Beach, CA, USA, pp. 291–301.

Boure, J.A., Bergles, A.E., Tong, L.S., 1973. Review of two-phase flow instability. Nucl. Eng. Des. 25, 165–192.

Bouton, C., Pagès, G., 1993. Self-organization and a. s. convergence of the Kohonen algorithm with distributed stimuli. Stoch. Process. Their Appl. 47, 249–274.

Cavalaris, J.G., 1998. Optimizing the Data Transmission Protocols for Remote Interactive Microscopy. Massachusetts Institute of Technology.

Chen, W., Pratt, W.K., 1984. Scene adaptive coder. IEEE Trans. Commun. Com. 32, 225–232.

Chiang, J.-H., Aritomi, M., Inoue, R., Mori, M., 1994. Thermo-hydraulics during start-up in natural circulation boiling water reactors. Nucl. Eng. Des. 146, 241–252.

de Mesquita, R.N., Penha, R.M.L., Masotti, P.H.F., Sabundjian, G., Andrade, D.A., Umbehaun, P.E., Torres, W.M., Conti, T.N., Macedo, L.A., 2010. Two-phase flow patterns recognition and parameters estimation trhough natural circulation test loop image analysis. In: 7th ECI International Conference on Boiling Heat Transfer. Florianopolis, Santa Catarina, Brazil.

de Mesquita, R.N., Ting, D.K., Cabral, E.L.L., Upadhyaya, B.R., 2004. Classification of steam generator tube defects for real-time applications using Eddy current test data and self-organizing maps. Real Time Syst. 27, 49–70.

de Mesquita, R.N., Masotti, P.H.F., Penha, R.M.L., Andrade, D.A., Sabundjian, G., Torres, W.M., Macedo, L.A., 2012. Classification of natural circulation two-phase flow patterns using fuzzy inference on image analysis. Nucl. Eng. Des. 250, 592–599. http://dx.doi.org/10.1016/j.nucengdes.2012.06.014.

Delhaye, J.M., Giot, M., Riethmuller, M.L., 1981. Thermohydraulics of Two-Phase Systems for Industrial Design and Nuclear Engineering. McGraw-Hill, New York, USA.

Emami, N., 2016. Application of image editing software for forensic detection of image. J. Fundam. Appl. Sci. 8, 1300–1307. http://dx.doi.org/10.4314/jfas.v8i2s.277.

Fu, Y., Liu, Y., 2016. Development of a robust image processing technique for bubbly flow measurement in a narrow rectangular channel. Int. J. Multiph. Flow 84, 217–228. http://dx.doi.org/10.1016/j.ijmultiphaseflow.2016.04.011.

Hsieh, C., Wang, S., Pan, C., 1997. Dynamic visualization of two-phase patterns in a natural circulation loop. Int J. Multiph. Flow 23, 1147–1170.

Kattan, N., Thome, J., Favrat, D., 1998. Flow boiling in horizontal tubes: Part 1 – development of a diabatic two-phase flow pattern map. J. Heat Transf. ASME 120, 140–147.

Kirouac, G.J., Trabold, T.A., Vassallo, P.F., Moore, W.E., Kumar, R., 1999. Instrumentation development in two-phase flow. Exp. Therm. Fluid Sci. 20, 79–93.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proc. of IJCAI'95 1137–1145. https://doi.org/10.1067/mod.2000.109031.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. 69, 59–69.

Kohonen, T., 2001. Self-Organizing Maps. Springer-Verlag, New York, USA.

LCIS, L. of C. and I.S., 2011. Som Toolbox.

Ledinegg, M., 1938. Instability of flow during natural and forced circulation. Die Warme 48, 891–898.

Mathworks Inc., T., 2015. Matlab.

Maurus, R., Ilchenko, V., Sattelmayer, T., 2002. Study of the bubble characteristics and the local void fraction in subcooled flow boiling using digital imaging and analysing techniques. Exp. Therm. Fluid Sci. 26, 147–155.

Nayak, A.K., Sinha, R.K., 2007. Role of passive systems in advanced reactors. Prog. Nucl. Energy 49, 486–498.

Nayak, A.K., Dubey, P., Chavan, D.N., Vijayan, P.K., 2007. Study on the stability behaviour of two-phase natural circulation systems using a four-equation drift flux model. Nucl. Eng. Des. 237, 386–398. http://dx.doi.org/10.1016/j.nucengdes.2006.05.009.

Pouryoussefi, S.M., Zhang, Y., 2015. Identification of two-phase water-air flow patterns in a vertical pipe using fuzzy logic and genetic algorithm. Appl. Therm. Eng. 85, 195–206. http://dx.doi.org/10.1016/j.applthermaleng.2015.04.006.

Pratt, W.K., 1972. Generalized Wiener Filtering Computation. IEEE Trans. Comput. C 636–641.

Ruspini, L.C., Marcel, C.P., Clausse, A., 2014. International journal of heat and mass transfer two-phase flow instabilities: a review. Int. J. Heat Mass Transf. 71, 521–548. http://dx.doi.org/10.1016/j.ijheatmasstransfer.2013.12.047.

Seleghim Jr, P., Hervieu, E., 1998. Direct imaging of two-phase flows by electrical impedance measurements. Meas. Sci. Technol. 9, 1492–1500.

Shamoun, B., Beshbeeshy, M. El, Bonazza, R., 1999. Light extinction technique for void fraction measurements in bubbly flow. Exp. Fluids 26, 16–26.

Shi, B., Zhang, Y., 2015. Phase discrimination and a high accuracy algorithm for PIV image processing of particle – fluid two-phase fl ow inside high-speed rotating centrifugal slurry pump. Flow Meas. Instrum. 45, 93–104. http://dx.doi.org/10.1016/j.flowmeasinst.2015.05.002.

Villasenor, J.D., 1993. Full-frame compression of tomographic images using the discrete Fourier transform. In: Data Compression Conference. DCC '93. IEEE, USA, pp. 195–203.

Wilhelm, P., Haynor, D.R., Kim, Y., Riskin, E.A., 1991. Lossy image compression for digital medical imaging systems. Opt. Eng. 30, 1479–1485.