The 2013 World Congress on
Advances in Structural Engineering and Mechanics (ASEM13)
Jeju, Korea, September 8-12, 2013

SHM with wave diffraction patterns and probabilistic decision tree based multi-class SVM

*Mechbal Nazih[1] and Uribe Juan Sebastian[2]

[1,2] *Processes and Engineering in Mechanics and Materials Laboratory*
PIMM, UMR CNRS 8006, ENSAM
151, Boulevard de l'Hôpital 75013, Paris, France
[1] *nazih.mechbal@ensam.eu*

ABSTRACT

In this work, a new process of damage and impact detection on structures with piezoelectric active elements using acoustical wave diffraction patterns and Probabilistic Decision Tree (PDT) tree Support Vector Machine (SVM) architecture is presented. This active SHM approach uses permanently emission of selected non-resonant Lamb waves into the structures and monitors a damage index (DI) relying on the recognition of amplitude disturbed diffraction pattern (ADDP). Based on this ADDP, a detection and localization approach is proposed. It exploits the measurements to train an original SVM clustering algorithm utilizing a specialized binary decision tree (SVM-PDT) producing a posteriori probabilities of damage localization in a multi-class context. The proposed SHM procedure is illustrated on actual plates.

1. INTRODUCTION

Advanced structures with self-capabilities have been intensively studied over these last four decades. These structures also called *smart structures* are emerging as a promoting way to improve the intrinsic and extrinsic characteristics of a structure. These intense research efforts represented the application of multidisciplinary monitoring and control methods to assess smart structures, seeking essentially safe operation, useful life extension, generally based on evaluation of their operational status, and eventually applying active control methods to ensure the required performance.

Structural health monitoring (SHM) is an emerging technology to automate the inspection process to assess and evaluate the health condition of structures in real-time or at specified time intervals. SHM systems for smart structures may automatically

[1] Professor
[2] Graduate student

process data, assess structural condition, and signal the need for human intervention (Worden, et al., 2007). SHM technology involves multidisciplinary fields ranging from material, structure, signal processing, data mining, fracture mechanics, fatigue life analysis and more. It aims to detect, localize and evaluate the severity of damages. Recent surveys have shown that even reluctant industry areas are now convinced that SHM is the key technology to enable the transition from schedule-driven maintenance to condition-based maintenance (Chang, 2011).

To perform damage monitoring, a variety of techniques have been developed. They could be classified into: passive or active and then into two sub categories: global or local. Passive SHM are techniques that only "listen" to the structure and infer its state by monitoring over time passive sensors. Although these methods have proved useful, their reliability could be increased by making them interacting with the structure. In this case, we speak about active SHM methods because actuators are used to interrogate the structure to enhance damage detection and localization. The extraction of damage-sensitive features from measurements is a process that can encompass either physical modelling or data-driven methodologies to SHM, but is most powerful when it underpins the latter and is based on pattern recognition (Farrar & Worden, 2012).

Ultrasonic inspection is well-established and widely used in several industrial domains as an efficient NDE technique. In fact, SHM can be viewed as embedded NDE. However, acoustic emission (AE) is a passive SHM method that could be prone to contamination by environmental noise (Su & Ye, 2009). One way to overcome this problem is to use an active method that generates specific signal processing to activate elastic waves that are greater than half the size of their wavelength in spite of the presence of environmental noises. In these active methods, two approaches could be considered: traveling waves or standing waves. And the interrogating is done on demand or continuously.
A large number of waves based techniques exist for SHM. These techniques exploit surface acoustic waves (SAW) or guided waves in plates, shells, or tube-like structures, to localize acoustic sources or damage. In recent years, new SHM methods using acoustic waves have been developed. They may be based on thermosonics (Barden, et al., 2007) or on a low power guided wave system (Kim, et al., 2009). An interesting research presents the use of a probability-based imaging algorithm for damage detection with Lamb wave signals (Lu, et al., 2009).

An approach through correlation technique has been presented recently (Leblanc, et al., 2007). This technique is applied to localize the formation of cracks or AE (acoustic emission) source. The impulse responses are acquired along a contour line that runs along the area of inspection. Internal impulse responses to the area are calculated using a wave superposition method. The detected acoustic waves mainly correspond to bending waves (A0 mode) through the assembly of non-symmetric piezoelectric plates. However, this process based on an absorption technique cannot address some particular damages such as static or quasi-static impacts that create more diffraction effect of the propagating acoustic waves than absorption.

In this work, we propose a correlation technique that relay on wave diffraction patterns recognition. It is an active damage monitoring process based on the amplitude disturbed diffraction pattern (ADDP) phenomenon of permanently emission of selected non-resonant Lamb waves. It assesses the disturbances that damage brings to the acoustic wave propagation in a solid. These disturbances depend on the damage position and the frequency of excited signal. With analysis and a calibration procedure of this variation, we can locate precisely the damage position. This process has been successfully used as a multi-touch sensing approach to tactile sensing (Liu, et al., 2009; Liu, et al., 2010). It has been diverted from their initial use to be applied to damage monitoring (Liu, et al., 2011).

The output of an SHM process should give information on the type, localization and the severity of damage. Moreover, it should give a fixed number of the more relevant damage features classified in a likelihood sense. In a pattern recognition approach, it has to be seen as probabilistic multi-class classification problem where each class represents a damage feature's (localization, type, ...). We propose here to describe an original classification technique, the Probabilistic Decision Tree (PDT) producing a posteriori probabilities in a multi-class context. It is based on a Binary Decision Tree (BDT) with Probabilistic Support Vector Machine classifier (PSVM). At each node of the tree, a bi-class SVM along with a sigmoid function are trained to give a probabilistic classification output. For each branch, the outputs of all the nodes composing the branch are combined to lead to a complete evaluation of the probability when reaching the final leaf (representing the class associated with the branch). Formally, we are interested in solving a multi-class data classification problem in a manner that produces confidence probabilities associated with each damage feature.

Actually, there exist two main types of classifiers: hard and soft (Liu, et al., 2011; Wahba, 2002). Hard classifiers, such as support vector machine (SVM) and all the associated multi-class techniques, build a frontier between classes. They only label new unknown points with the class associated to the side of the frontier in which they fall, without giving any idea of the certitude of the decision or the degree of membership to that class. These classifiers are very appealing, because in general they tend to give very accurate predictions. On the opposite, soft classifiers like Logistic Regression (LR) (David W. Hosmer & Lemeshow, 2004) are able to build probability estimations for the belonging to all the classes, and then with this information they choose the most likely class. We are thus interested, in a multi-class context, by the probabilities estimation offered by soft classifiers while keeping the hard classifiers proved performances (Platt, 2000). To reach this objective, hard classifiers need to be first adapted to a multi-class context. Support Vector Machines (SVM) are a powerful tool for data classification (Weston & Watkins, 1998). Unfortunately, they were originally designed for bi-class decision problems and their extension to multi-class problems is not straightforward and is still an on-going research issue (Hsu & Lin, 2002). Classic SVM multi-class approaches such as *"one-against-one"* (OvO) (Friedman, 1996), *"one-against-all"* (OvA) (Vapnik, 1998) or Diagram Acyclic Graph (DAG) (Platt, et al., 2000) have shown adequate results but don’t take into account the structure and the distribution of the data when separating the classes. To overcome this drawback, Madzarov (Madzarov, et al., 2009) came up with a simple and intuitive approach based on building a binary decision tree. By selecting specific features, such as the distance between gravity centers of the different classes, an automatic graph is generated where at each node a bi-class SVM is trained. However, such multi-class hard classifiers only provide one predicted class without any associated score indicating the confidence of the classification. In this bi-class context, Platt (Platt, 2000) proposed a method for extracting probabilities $p(class|input)$ from SVM outputs to be used for classification post-processing. The approach consists in training the parameters of a sigmoid function to map the SVM outputs into probabilities. The underlying idea of Probabilistic SVM classifier (PSVM) is that as the distance from an example to the frontier is larger, the example is closer to that class, which implies that the example will very likely belong to that class. Adapting Platt's method to a multi-class context, it is thus, in principle, possible to build the confidence index that we need while keeping the demonstrated performances of hard classifiers.

Based on Madzarov (Madzarov, et al., 2009) and Platt (Platt, 2000) algorithms, we present Probabilistic Decision Trees (PDT) as an original approach to the multi-class probabilistic classification problem. The proposed PDT algorithm takes advantage of the decision tree architecture and of the classification posterior probability provided by PSVM. The PDT will provide fast classification (logarithmic complexity) along with associated posterior probabilities $p(class|input)$. At each node of the PDT, SVM classification associated with a sigmoid function is performed to estimate the probability of membership to each sub-group. A probability function is then built for each leaf, by following the path that the PDT has generated for it.

The remit of this paper is to present an original active data-driven SHM approach. It is based on acoustical wave diffraction patterns as damage-sensitive features and a Probabilistic Decision Tree (PDT) tree Support Vector Machine (SVM) architecture for damage indicator. The approach will follow sequential process depicted in Fig. 1.

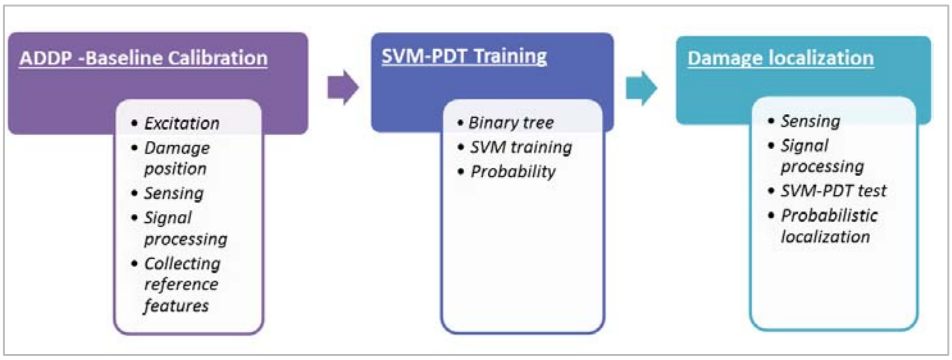


Fig. 1: ADDP–SVM-PDT probabilistic damage localization process

2. LAMB WAVES ILLUMINATION

Consider a structure like plate with Lamb wave transmitters and receivers on its edges, as the most energy carried out by a Lamb wave is the shear wave energy, it follows, that in case of static damage this energy will be attenuated which will provide position associated information for localization. Indeed, once stable wave interference is established on the structure, it will be monitored by wave receivers (Fig. 2). In the case of damage, diffraction signals will be observed in real time by the receivers. If the diffraction signal has a bijective relation with the damage's property, e.g. position, we can then use a pattern recognition process to identify the damage.

This observation supposes that the plate is permanently excited with acoustic waves. We have proposed a nondestructive active flaw monitoring method using non-resonance acoustic waves composed of several (> 20) frequency components. The monitoring process relays on the identification of waves diffraction patterns. This process is also called an *Amplitude Disturbed Diffraction Pattern* (ADDP) process.
Exciting the structure with non-resonance frequencies permit also to overcome the problems associated with the instability of resonance patterns. In addition, since there is no more resonance materialized on the surface of the object, to help the localization of flaw or a touch on the plate, this process suggests to replace resonance patterns with figures of illumination. In fact, the method presented in this study proposes to create pattern which is not dependent on the resonance modes with high selectivity, but on the way how waves propagate through the plate. For more details and analytical consideration on this approach please refer to Liu (2010).

As suited previously, the Lamb wave pattern method requires first a calibration process. This process is described in Fig. 2. Consider a rectangular structure with wave transmitters and receivers, as shown in Fig. **3**. For each possible damage position, discrete points (x, y) are defined *a priori* on the surface. An acoustic signal measured in the calibration step associated with one discrete point, is considered as one pattern. A calibrated impact on the object in the localization step will give a diffracted signal of the propagating Lamb waves. The localization involves a classification process.

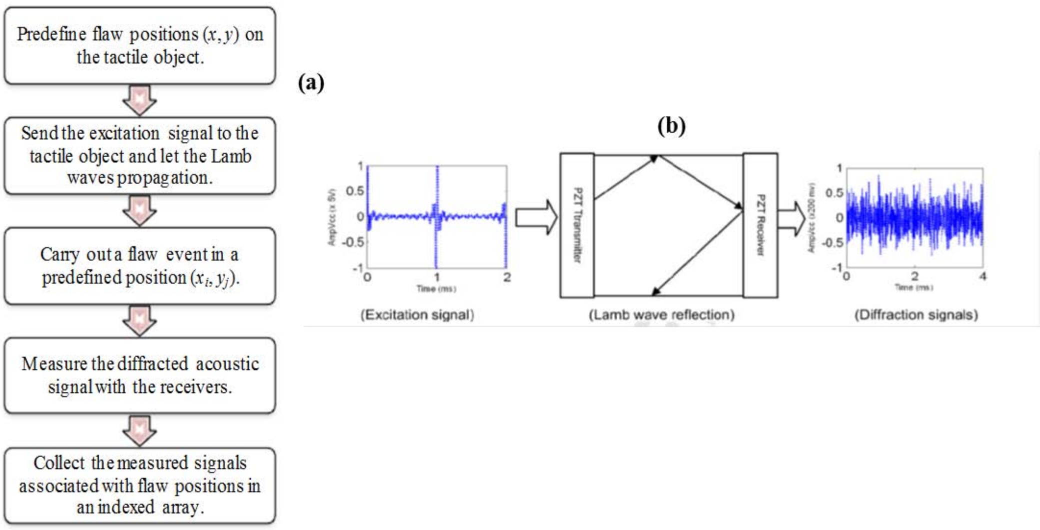


Fig. 2: (a) Calibration steps. (b) Signal flux

If the emitted signal is given as $e(t)$, the received acoustic signal $r(t)$ can be expressed as a function of emitted signal, the object geometry G, the flaw position P and flaw surface S:

$$r(t) = F(e(t), G, P, S) \tag{1}$$

In a training step, the sensible surface of the object is divided into an array of predefined "flaw points". Predefined flaw points are sequentially realized to stock a reference matrix. The area and nature of flaw is identical as a first approximation. The Fourier transform, $R_{i,j}$, of received signal at each position $P_{i,j}$ can be expressed by:

$$R_{i,j}(f) = F(E(f), G, P_{i,j}, S) \tag{2}$$

If the flaw effect is minor comparing with the whole object, the $E(f)$ and G can be considered as constants. As we maintain the flaw surface, the variation of S is negligible, so we consider that the received signal $r(t)$ is only function of contact position P.

$$R_{i,j}(f) = F(P_{i,j}, f) \tag{3}$$

Once we have an array of reference received signals, we can use it for monitoring the damage by generating damage index and feeding the SVM-PDT classifier.

In the case of impact detection, a test bench has been designed (Fig. 7). It is a copper plate with dimension 75 mm × 100 mm and active elements (PZT,type Pz27) bonded to the edges.

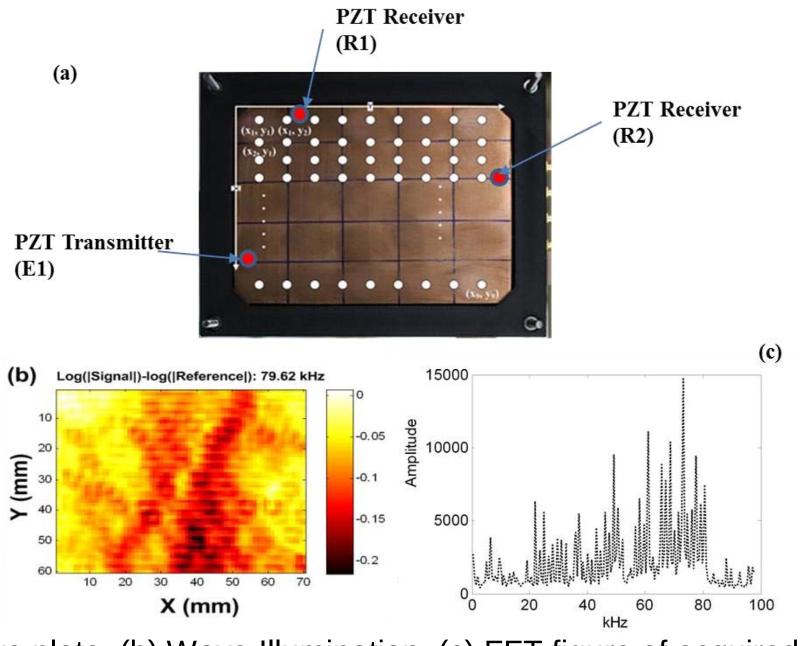


Fig. 3**:** (a) Active plate. (b) Wave Illumination. (c) FFT figure of acquired signals. Quasi-static impact at the position (x1, y1), by PZT Receiver 1

This method enhances the active acoustic plate featuring a process that continuously generates Lamb wave packets, 1000 per second. This monitoring technology is said "active" as regards Lamb wave generation, meaning excitation signals are fundamental to its reliability.

4. PROBABILISTIC CLASSIFICATION

4.1 Probabilistic Support Vector Machines

Classic support vector machines (SVM) have proved to be a very effective classification method (Lee, et al., 2001). They are binary linear classification techniques which search for the hyper plane (in the hyperspace of attributes) that separates two classes in a training set. This hyper plane is found by maximizing the so-called margin, which is the distance from the hyper plane to the closest points, denoted support vectors. A common variant of classic SVM, is called *soft margin* and it consists of admitting some misclassified points in the training set for preventing the over fitting problem. Although, we want to avoid too many points being misclassified, thus we impose a penalty C that will penalize every misclassified example. C can take values in the range $0 < C \leq \infty$. A high value of C means a strict classifier that doesn't admit many misclassified points. On the opposite, a small C means a very flexible classifier. Formally, we have a training set $\{ (\boldsymbol{x_1}, \boldsymbol{y_1}), \dots, (\boldsymbol{x_n}, \boldsymbol{y_n})\}$, where every point $\boldsymbol{x_i} = (\boldsymbol{x_{i1}}, \dots, \boldsymbol{x_{im}})$ has m attributes and one of the two possible labels $\boldsymbol{y_i} = \{-1, 1\}$. A soft margin SVM classifier will label a new unknown point $\boldsymbol{x_t}$ according to the decision function:

$$y(\mathbf{x}_t) = sign((\boldsymbol{w}.\boldsymbol{x}_t) + w_0) \tag{5}$$

where w and b are the hyper plane parameters obtained from the minimization of the cost function in Equation 7 on the training set.

In SVM, kernels are used to project the data into a virtual space where it might be easier to separate them (Vapnik, 1998). The main advantage of kernel functions is that the only operation needed to be defined in the new virtual space is the inner product $\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle$. Several kernel functions are used in this work we manly use the Gaussian kernel (Chang & Lin, 2011).

Applying a kernel function, the soft margin and the Wolfe dual formulation the SVM problem is presented as:

$$y(\mathbf{x}_t) = sign(f(\boldsymbol{x_t})) \tag{6}$$

where $f(\boldsymbol{x})$ is defined as:

$$f(\boldsymbol{x}) = \sum_{i=1}^{m} \alpha_i u_i \kappa(\boldsymbol{x}_i, \boldsymbol{x}) + w_0 \tag{7}$$

the values of $\alpha_{\rm i}$ and u_i are found solving the following constrained optimization problem:

$$\begin{cases} \max_{\alpha_i} \left[\sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j u_i u_j \kappa(\boldsymbol{x_i}, \boldsymbol{x_j}) \right] \\ 0 \leq \alpha_i \leq C, \; i = 1, \dots, m \\ \sum_{i=1}^{m} \alpha_i u_i = 0 \end{cases} \tag{8}$$

In the PDT approach, and for its use in the car diagnostic classification, we choose to use the Gaussian kernel. It only has σ^2 as parameter. A small value of σ^2 will lead to curved hyper plans and a high value will force the hyper plans to be straighter. In

(Keerthi & Lin, 2003) it is showed that from some combinations of the hyperparameters (C,σ^2), the Gaussian kernel tends towards the linear kernel, which makes the Gaussian kernel the most general method and one that will work for a large range of datasets. The hyper parameters (C,σ^2) have to be optimized for every classification problem. In (Keerthi & Lin, 2003) an effective technique that we implemented is proposed.
As stated before, SVM produce a value that is not a probability. Indeed, SVM only give a class prediction output that will be either +1 or -1. In order to extract associated probabilities from SVM outputs several approaches have been proposed (Platt, 2000; Vapnik, 1998; Hastie & Tibshirani, 1998). We will focus on Platt's approach (Platt, 2000). Platt (Platt, 2000) proposed a technique that has been largely used in the literature. He builds a sigmoid function between the outputs $f(\boldsymbol{x})$ of the SVM and the probability of membership $p(y=i|\boldsymbol{x})$ to a class i, given the attributes of $\boldsymbol{x}$. A simple bi-class example is shown in Fig. 4
The sigmoid will have the following parametric expression:

$$p\,(y\; = 1|f(\boldsymbol{x})) = \frac{1}{1+e^{af(x)+b}} \tag{5}$$

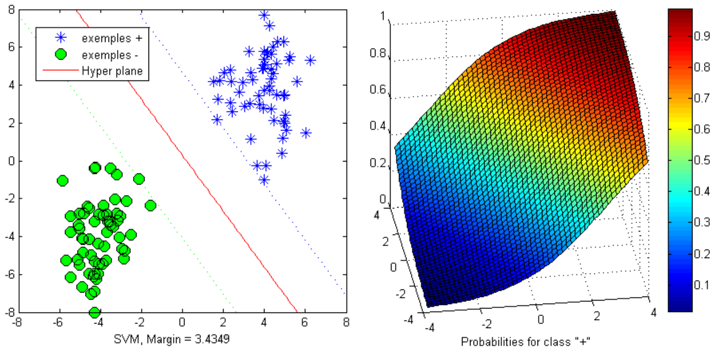


Fig. 4: SVM and probability estimation for a 2D binary problem

where a and b are parameters computed from the minimization of the negative log-likelihood function (Platt, 2000)

$$-\sum_i t_i\,\log p(f(\boldsymbol{x_i}))\; + (1-t_i\,)\log(1-p(f(\boldsymbol{x_i})) \tag{6}$$

and t_i is the new label of the classes. 1 becomes t_+ and -1 becomes t_-. This relabeling procedure is conducted so that the sigmoid fit will be softer. These new labels are computed using the expressions:

$$t_+ = \frac{N_+ + 1}{N_+ + 2} \qquad t_- = \frac{1}{N_- + 2} \tag{7}$$

where N_+ and N_- are the number of points that belong to class 1 and class 2 respectively.
The PSVM as proposed by Platt (Platt, 2000) uses first a SVM classifier that has to be trained with a training set s_1. Then the sigmoid parameters (a,b) have to be found. To do so, it is recommended (Platt, 2000) to use a second training set s_2. The sigmoid fit requires the inputs $f(\boldsymbol{x})$ and the labels t_i. We use s_2 as a test set for the classifier. Thus we obtain the output values $f(s_2)$. Knowing the labels of s_2 (how many positives

ones and how many negatives ones) we can calculate t_+ and t_- for this sub-training set. With values $(f(s_2), t_i(s_2))$ we then find (a, b) by optimizing Eq. (9) using a Newton's method with backtracking, as proposed by (Lin, et al., 2007). After these two steps, we have an a posteriori probability estimator to the bi-class problem. We can then classify new points and give them an associated probability of belonging to class 1 or 2.

4.2 From bi-class to multi-class problems

SVM were originally designed for bi-class classification problems, the passage to multi-class problems is still an on-going research area (Hsu & Lin, 2002). There are two major approaches to solve this type of problems.

The first and the one that could be more intuitive consist in formulating a cost function with Q (the number of classes in our problem) hyperplanes (Weston & Watkins, 1998).

$$\begin{cases} \min_{\boldsymbol{w},\xi} \dfrac{1}{2}\sum_{m=1}^{Q}(\boldsymbol{w_m}.\boldsymbol{w_m}) + C\sum_{i=1}^{l}\sum_{m\neq y_i}\xi_i^m \\ (\boldsymbol{w_{y_i}}.\boldsymbol{x_i}) + b_{y_i} \geq (\boldsymbol{w_m}.\boldsymbol{x_i}) + b_m + 2 - \xi_i^m \\ \xi_i^m \geq 0, \quad i = 1, \dots, l \quad m \in \{1, \dots, Q\} \setminus y_i \end{cases} \tag{8}$$

For this formulation, the decision is given by

$$f(\boldsymbol{x}) = \max_k(\boldsymbol{w_i}.\boldsymbol{x_i}) + b_i \ , \quad i = 1, \dots, Q \tag{9}$$

This method suffers from the problem that an optimization with so many variables is more difficult to solve for the algorithm, can give slower results and even in some occasions may not converge (Hsu & Lin, 2002).

Another approach is to divide the multi-class problem in several binary sub-problems. There are numerous methods that do the division. The most popular are *One against one* (Friedman, 1996), *One against all* (Vapnik, 1998) , *Diagram Acyclic Graph (DAG)* (Platt, et al., 2000) *and Binary decision Trees (BDT)* (Madzarov, et al., 2009). For sake of shortness, we will only focus on the BDT method, because it is the one we will use.

In (Madzarov, et al., 2009), they proposed to build a binary tree in which at every node the remaining classes separated in two subgroups g_1 and g_2. A SVM classifier decides to which subgroup the new point belongs, so in which direction to move. In order to build the tree (the first step in the classification procedure) a clustering algorithm divides all the Q classes into g_1 and g_2. The algorithm calculates the gravity centers of all classes, the two classes with the biggest Euclidean distance from each other are assigned to each of the sub-groups. Then the algorithm checks the closest class to one of the sub-groups, this class is assigned to that sub-group, their gravity center is recalculated with the new points that have just been added. This is repeated until all classes have been assigned to one of the groups. An example is illustrated by of Fig. 5. For each sub-group the clustering algorithm is repeated until there is no more than one class per sub-group. Those sub-groups will be called the leaves of the tree and a point that falls there will be assigned with the class of the leaf.

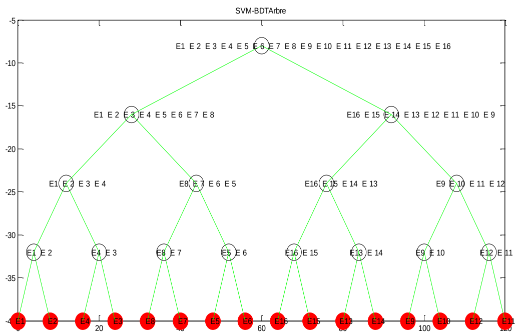


Fig. 5: Illustration of SVM-BDT

It is important to note that BDT testing time is smaller than other methods because the depth of the decision tree is of order $\log_2 Q$ since at every level the tree eliminates approximately half of the remaining classes. The testing time is an important characteristic that must be taken into account when choosing a classifier.

4.2 Probabilistic decision trees (PDT)

In this paper, we propose to build a binary decision tree following the idea introduced by Madzarov (Madzarov, et al., 2009), but instead of using a simple SVM classifier in each node, we use a SVM classifier associated with a sigmoid function (PSVM) to estimate the probability of membership to each sub-group in the node, as shown in Fig. 6. The tree may be built using different criteria, for example the Euclidian distance between the gravity centers, the margin obtained by pairwise SVM (Chalasani, et al., 2007) or some physical or functionalities criteria (mechanical, electrical, etc.). This is very interesting because in this way we introduce previous knowledge from an expert that might help the classification task. We can then build a probability function for each leaf, knowing the path that a point has to follow to reach it.

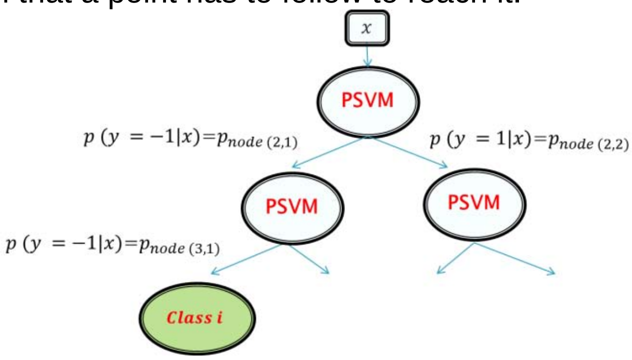


Fig. 6: Example of a probability decision tree

Note that there is only one way to get to a leaf, so the probability functions are unique for a trained tree.

$$p(y = i|x) = \prod_{h=1}^{leaf} p_{node\,(h,l)} \tag{9}$$

h is the level of the tree and $h = 1$ is the root node. The previous expression states that the probability of membership of an element to the class i is calculated as the product of the probabilities of the decisions taken in all the nodes visited until arriving to the leaf. By $node(h, l)$ we mean the l node in the h level. Once the tree is built we will also have the Q probability functions, one for each class. When classifying future unknowing cases we will just have to evaluate the Q analytical functions and then chose the class with the highest score. However, we don't have to settle with just one predicted class. One of the most interesting things of the PDT is that we can have more than one prediction for one subject. We can have a list of all the possible classes, ordered after their plausibility, which is measured with the probabilities estimation. So, instead of having one predicted class like classic SVM multi-class methods, we will have several options.

5. DAMAGE LOCALIZATION ON AN ACTIVE COMPOSITE PLATE

5.1 Experimental setup

The proposed damage localization method is applied to a free-free composite plate $(300 \times 400 \times 0.4mm)$ with 4 PZT transducers. The experimental setup is described in Fig. 7.

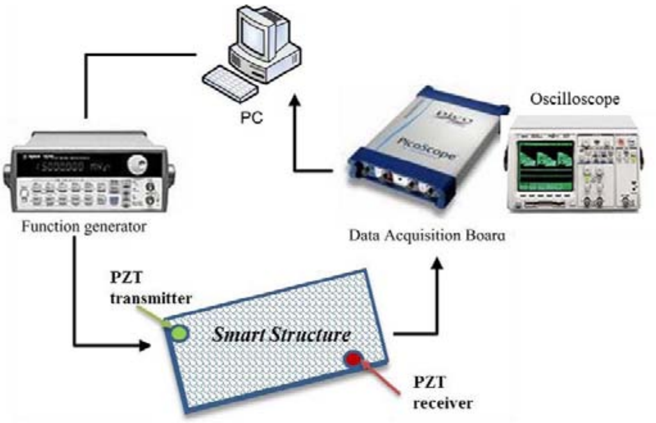
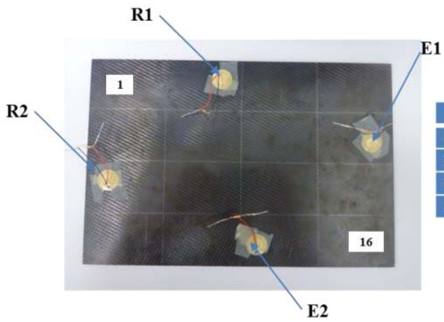


Fig. 7: The smart structure and the test bench

The plate is divided into 16 zones (Fig. 8). To simulate damage, a calibrated mass has been positioned on each zone.



	Col 1	Col 2	Col 3	Col 4
Row 1	1	2	3	4
Row 2	5	6	7	8
Row 3	9	10	11	12
Row4	13	14	15	16

Fig. 8: Composite plate and damage zones

Lamb waves at different frequencies have different propagating velocities and ratios between the normal and the shear components, so they have different diffraction behaviors when a finger touches the plate. A sinusoidal excitation signal could provide single amplitude change information produced by damage. Adequate number of different excitation frequencies is required to distinguish the damages at numerous predefined positions, which would have same or similar modified amplitude at same frequency. In this paper, we propose an excitation signal composed of 31 frequency components distributed from 20 kHz to 100 kHz.

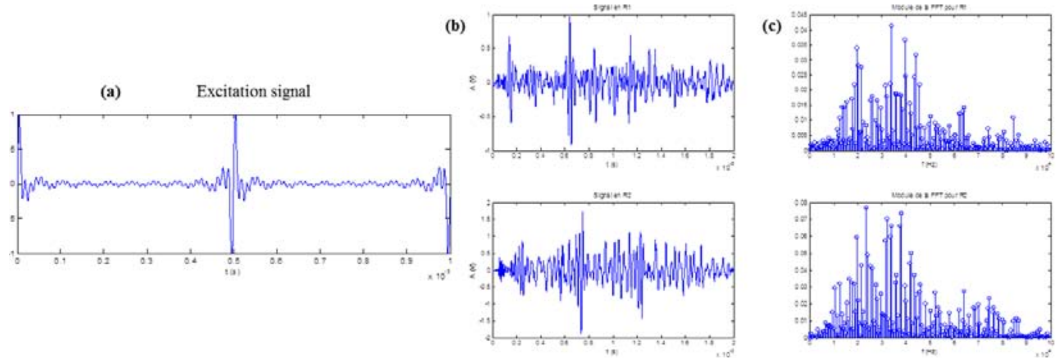


Fig. 9: (a) Excitation signal. (b) Time response. (c) Frequency response.

5.2. Calibration and SVM-PDT classification

A PSVM-PDT classifier has been trained using a dataset composed by the 96 experiences made on the plate (6 repetitions for each position). As attributes we chose the FFT (Fast Fourier Transform) response of the signal received by each sensor R1 and R2. In order to extract only the most significant information, we use only the first 200 frequencies of the spectrum.
Given that each sensor perceives different parts of the information, we have concatenated both signals, obtaining a vector containing 400 attributes.

$$x_i = [\, f_1(R_1), f_2(R_1), \dots, f_{200}(R_1), f_1(R_2), \dots, f_{200}(R_2)] \quad i = \{1,96\}$$

The labels will be the 16 zones, $y_i = \{1,16\}$, in which the plate is divided, as shown in Fig. 8

Once trained, the PDT has Q probability functions, being Q the number of zones in which we have split our plate. The damage position can be visualized on a figure representing the normalized distances with a grey amplitude scale. In Fig. 10, the probability distribution over the plate is shown for a test that was made adding an artificial damage in zone 16. This example was used in the learning database. It's very interesting to see how the zone is well isolated, the neighbors also have a probability of being the damaged zone, but a smaller one.

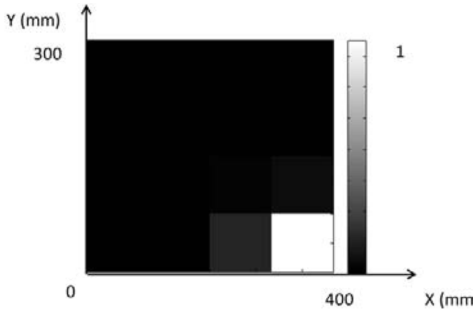


Fig. 10: Probability distribution for a damage located in zone 16

To test the effectiveness of the method, we give to the classifier 8 "unknown" points to classify. These points are other experiences performed by introducing the damage anywhere on the plate (Fig. 11). For these points, the classifier will only have their FFT information, as above

$$x_{p_i} = [\, f_1(R_1), f_2(R_1), \dots, f_{200}(R_1), f_1(R_2), \dots, f_{200}(R_2)] \quad i = \{1{,}8\}$$

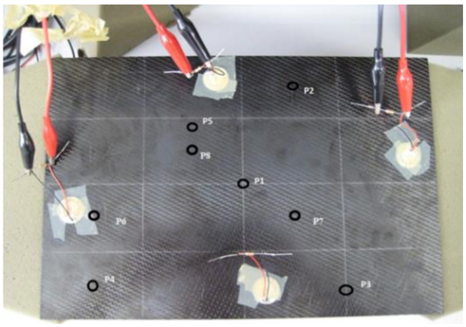


Fig. 11: Test damage location

The predicted damages localization (P6 and P3) and the their associated probability are given in Fig. 12 and Fig. 13.
For test P3 the damage is located in the last row, between the third and fourth colon. The classifier has located the damage in the right position. We can see that it also gives a small probability to the zones around, since the damage was not positioned in the middle of the zone.

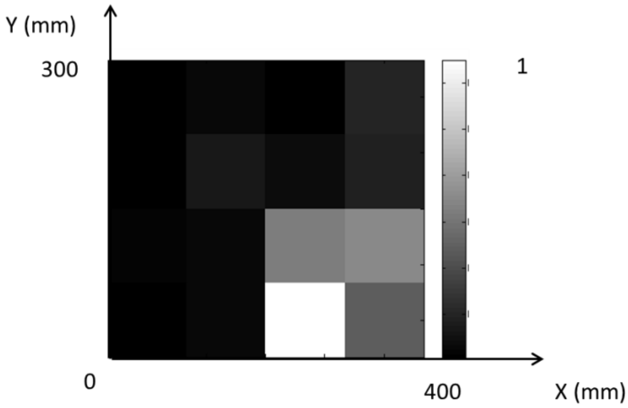


Fig. 12: Probability distribution for the test P3

For test P6 the damage is located in the third row, first colon. The classifier has located the damage in the right position. There are some zones, in particular on the last colon, that have small probabilities too. These results could be expected due to wave reflections. To ovoid these ambiguity, more frequencies should be used in the excitation signal.

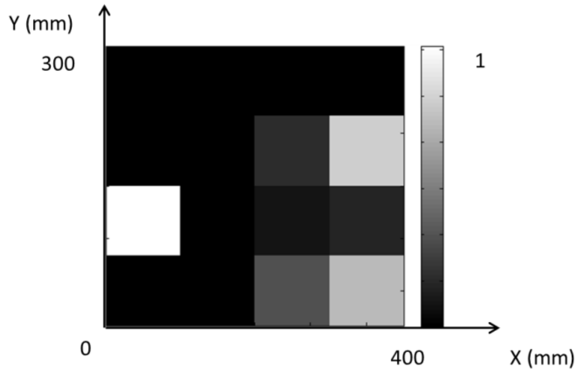


Fig. 13: Probability distribution for the test P6

3. CONCLUSIONS

In this paper, we have proposed an original active data-driven SHM approach. It is based on acoustical wave diffraction patterns as damage-sensitive features and a Probabilistic Decision Tree (PDT) tree Support Vector Machine (SVM) architecture for damage indicator.

This active SHM scheme uses permanently emission of selected non-resonant Lamb waves into the structures and monitors a damage index (DI) relying on the recognition of amplitude disturbed diffraction pattern (ADDP). Based on this ADDP, a detection and localization approach is proposed. It exploits the measurements to train an original SVM clustering algorithm utilizing a specialized binary decision tree (SVM-PDT) producing a posteriori probabilities of damage localization in a multi-class context. The approach has been test experimentally on a composite plate, and results are promoting. To enhance the classification some new algorithmic developments are in progress based on the introduction of structure geometry's information during the construction the binary tree.

REFERENCES

Barden, T. et al., 2007. Detection of impact damage in CFRP composites by thermosonics. *Nondestructive Testing and Evaluation,* Volume 22(2), pp. 71-82.

Chalasani, T. K., Namboodiri, A. M. & Jawahar, C., 2007. *Support vector machine based hierachical classifieds for large class problems.* s.l., s.n.

Chang, C.-C. & Lin, C.-J., 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology,* Volume 2, pp. 27:1--27:27.

Chang, F.-K., 2011. *Structural Health Monitoring: Condition-based Maintenance.* Stanford, s.n.

David W. Hosmer, J. & Lemeshow, S., 2004. *Applied Logistic Regression.* s.l.:Wiley.

Farrar, C. R. & Worden, K., 2012. *Structural Health Monitoring: A Machine Learning Perspective.* s.l.:John Wiley & Sons.

Friedman, J., 1996. *Another approach to polychotomous classifcation,* Stanford: Stanford University, Department of Statistics.

Hastie, T. & Tibshirani, R., 1998. Classification by pairwise coupling. *The annals of*

statistics, 26(2), pp. 451-471.
Hsu, C.-W. & Lin, C.-J., 2002. A comparison of methods for multiclass support vector machines. *Neural Networks, IEEE Transactions on,* 13(2), pp. 415-425.
Keerthi, S. S. & Lin, C.-J., 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural computation,* 15(7), pp. 1667-1689.
Kim, J.-K., Zhou, D., Ha, D. & Inman, D., 2009. *A practical system approach for fully autonomous multi-dimensional structural health monitoring.* s.l., s.n.
Leblanc, A., Fink, M. & Ing, R., 2007. Senstive Materials. *Proceeding of the 8th International Workshop on Structural Heath Monitoring ,* Volume 1.
Lee, Y., Lin, Y. & Wahba, G., 2001. *Multicategory support vector machines (preliminary long abstract) Technical Report 1040,* Madison, WI: Department of Statistics, University of Wisconsin.
Lin, H.-T., Lin, C.-J. & Weng, R. C., 2007. A note on Platt's probabilistic outputs for support vector machines. *Mach. Learn.,* oct, 68(3), pp. 267-276.
Liu, Y., Mechbal, N. & Vergé, M., 2011. *Damage monitoring based on wave illumination of structures.* Stanford, USA, s.n.
Liu, Y. Nikolovski, J.P.; Mechbal, N.; Hafez M. & Vergé M., 2009. Tactile objects based on an amplitude disturbed diffraction pattern method. 95(25).
Liu, Y. Nikolovski, J.P.; Mechbal, N.; Hafez, M.& Vergé, M., 2010. An acoustic multi-touch sensing method using amplitude disturbed ultrasonic wave diffraction patterns. *Sensors and Actuators A,* Volume 162, pp. 394-399.
Liu, Y., Zhang, H. H. & Wu, Y., 2011. *Hard or soft classification? large-margin unified machines.* s.l.:Taylor & Francis.
Lu, Y. et al., 2009. *Lamb Wave Signals for Damage Detection Using a Probability-Based Imaging Algorithm.* s.l., s.n.
Madzarov, G., Gjorgjevikj, D., Chorbev, I. & others, 2009. A multi-class svm classifier utilizing binary decision tree. *Informatica,* 33(2), pp. 233-241.
Platt, J., 2000. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods.. In: *Advances in Large Margin Classifiers.* Cambridge: MIT Press.
Platt, J. C., Cristianini, N. & Shawe-Taylor, J., 2000. Large margin DAGs for multiclass classification. *Advances in neural information processing systems,* 12(3), pp. 547-553.
Su, Z. & Ye, L., 2009. *Identification of Damage Using Lamb Waves.* s.l.:Springer.
Vapnik, V., 1998. *Statistical learning theory. 1998.* s.l.:Wiley, New York.
Wahba, G., 2002. Soft and hard classification by reproducing kernel Hilbert space methods. *Proceedings of the National Academy of Sciences,* 99(26), pp. 16524-16530.
Weston, J. & Watkins, C., 1998. *Multi-class support vector machines Weston, J., & Watkins, C. (1998). Multi-class support vector machines. Technical Report CSD-TR-98-04,* London: Department of Computer Science, Royal Holloway, University of London.
Worden, K., Farrar, C. R., Manson, G. & Park, G., 2007. The fundamental axioms of structural health monitoring.. *Proceeding of the Royal Society*, Volume 463, pp. 1639-1664.