ESTIMETION OF THE CONCRETE CHARACTERISTICS USING PATTERN RECOGNITION METHODS

M.K. Titsias, D.I. Fotiadis and A. Likas

7 – 2001

Preprint, no 7 – 01 / 2001

Department of Computer Science
University of Ioannina
45110 Ioannina, Greece

ESTIMATION OF THE CONCRETE CHARACTERISTICS USING PATTERN RECOGNITION METHODS

M. K. Titsias, D. I. Fotiadis and A. Likas
Dept. of Computer Science,
University of Ioannina, GR-45110 Ioannina, Greece

1. SUMMARY

We address the problem of nondestructive analysis and testing of concrete. The proposed system employs acoustic emission and pattern recognition methods of the corresponding waveforms to identify the water to cement and sand to cement ratios of a given concrete cube. For waveforms obtained experimentally using standard size concrete cubes, we study several classification methods and provide comparative results.

2. INTRODUCTION

The nondestructive testing of concrete constitutes a problem which undoubtedly is of significant importance. Our objective is to extract chemical composition characteristics of concrete, such as the ratio of water to cement (W/C) and sand to cement (S/C), and finally determine the stiffness and the microstructure of a concrete sample. Some techniques of nondestructive testing of concrete presented in the literature are based on nuclear methods and the dielectric properties of concrete [1]. The relation of the chemical composition (W/C) of the concrete using ultrasound has been also addressed [1].

The method presented in this paper is a data driven approach since we use experimentally collected data (cases of concrete cubes with known geometrical and composition characteristics) in order to construct a pattern recognition system capable of predicting the composition of unknown concrete cubes. The system is able to determine the W/C and S/C ratio of unknown concrete cubes, classifying them into a predefined set of classes. The construction of the system is based on a database of waveforms [3] which is obtained from experiments carried out using acoustic emission. Several classification methods are employed to tackle the problem. More specifically, we considered two neural network techniques, based on the multilayer perceptron and the radial basis function network [4]. We applied also a classical non-parametric method, the k nearest neighbor as well as a statistical pattern recognition method, based on probability density estimation and the Bayes decision rule [4, 5]. In addition, we provide with comparative results, identify the most promising methods and present classification results which are very encouraging.

3. ACOUSTIC EMISSION IN CONCRETE CUBES

Acoustic emission can be used for nondestructive analysis and testing of objects and aims at detecting cracks or other defects of the object. In addition, it provides information concerning the response and behavior of a object under pressure, which is associated with the object strength and chemical composition.

In this work, we use a set of waveform features produced by experiments using the acoustic emission method [2] in concrete cubes. The complete setup of these experiments is described in [3]. A typical waveform and its corresponding features are shown in Fig. 1.

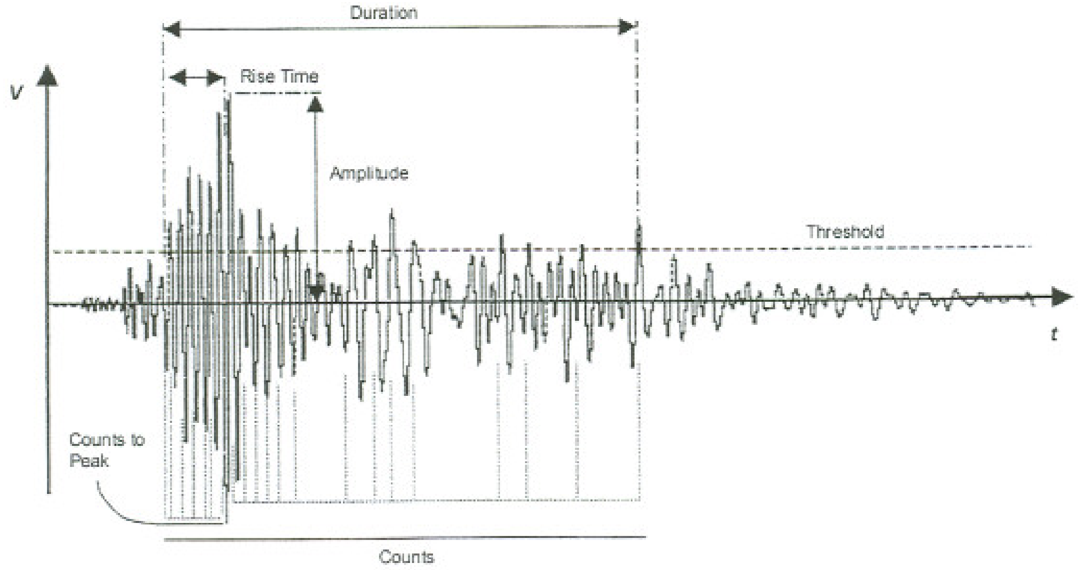

Figure 1: A typical waveform produced by acoustic emission and the corresponding features.

The terms used in Fig. 1 are explained below:

- *Threshold*: A specific signal amplitude used to separate the signal from noise. It is chosen such that to be slightly larger than the noise maximum amplitude.
- *Amplitude*: The maximum signal amplitude. This feature provides information about the intensity and the source distance.
- *Energy*: The area between the waveform and the time axis. It provides information about the intensity as well as the signal amplitude, the duration of signal and the signal its frequencies.
- *Counts*: The number of times the signal amplitude overruns the threshold. It depends on the signal intensity and frequencies.
- *Duration*: The time passed between the first the last overrun of the threshold. It provides information about the signal appropriateness in relation with the counts and the contained frequencies.
- *Rise time*: The time passed between the first threshold overrun and the occurrence of the maximum amplitude. It gives information about the intensity and the distance of the source. In addition, it can be used for noise detection.
- *Counts to peak*: The number of signal hits that overrun the threshold during rise time. It contains information, similar to the counts, for the signal part before the peak.

Note that the last six of the above terms correspond to features of the waveform and can be used by a pattern recognition method.

4. THE CLASSIFICATION PROBLEM

Our goal is to construct a classification system which will take as input concrete cube data (waveform features) of known age and provide with the ratio W/C and S/C, classifying the concrete cube into a predefined set of classes. The construction of the system is carried out using data driven methods (methods that use known data in order to train a system which can subsequently be used in unknown cases). The available waveforms have been obtained experimentally for cubes of four compositions concerning the ratio W/C and two concerning the ratio S/C, respectively (Table 1). In addition, the experiments have been performed on every concrete cube in four different ages: 1^{th}, 7^{th}, 28^{th} and 90^{th} day from production.

Table 1: The different W/C and S/C compositions of the concrete cubes and their names.

W/C \ S/C	0.50	0.55	0.60	0.65
3/1	A	B	C	D
4/1	E	F	G	H

Our classification approach is based on the fact that the total task can be partitioned into two subtasks: i) the estimation of the ratio W/C and ii) the estimation of the ratio S/C. Consequently, we could resolve the two problems independently. In such a case we can consider that the classification system consists of two classifiers which operate independently on the same input data in order to estimate W/C and S/C ratio. The estimation of the W/C ratio is a four-class problem and according to Table 1, data of compositions {A, E} correspond to the first class, {B, F} to the second class, {C, G} to the third class and {A, E} to the fourth class. Similarly, the estimation of the S/C ratio is a two-class problem and data of compositions {A, B, C, D} are from the first class, while data of compositions {E, F, G, H} are from the second class.

Now, each data point used for the construction of the classifier is essentially a feature vector with a label indicating its class. Such a feature vector can be constructed using all the waveform features described in Section 3. However, after a feature evaluation analysis applied to the total set of features, we found that the most suitable features for classification are: Energy, Counts and Duration.

5. CLASSIFICATION METHODS

We studied four classification methods. Two are based on neural networks methods, specifically on multilayer perceptron (MLP) and the radial basis function network (RBF). The third is the *k*-NN classifier and the fourth is a statistical pattern recognition technique which is based on probability density estimation and the Bayes decision rule. Next, we describe briefly the basic concepts of the methods and the way they have been used in our case.

The Multilayer perceprtron (MLP)

The multilayer perceptron or backpropagation network [4] constitutes the most widely used neural network architecture for classification problems. Given a K-class problem with d-dimensional input data the MLP architecture includes d inputs nodes, K outputs nodes and usually one or two hidden layers of non linear (sigmoidal) activation functions. For the solution of classification problems each class i is required to be encoded with a K-dimensional indicator vector which takes zero values everywhere except for the i^{th} component which is one. The MLP network is trained to map each input data vector into its indicator vector. The training process is based on the minimization of the mean square error function estimated for the set of training data. To classify an unknown input data point, network outputs are computed and the class corresponding to the largest output value is selected.

The most common method to train an MLP network is the backpropagation algorithm which is based on the gradient descent method. However, in our case where the amount of data is small and the number of network weights is not large, it is preferable to use more sophisticated optimization techniques such as the Newton methods. In our experiments we have used a Quasi-Newton method called BFGS which gave superior results compared to gradient descent and other methods. In addition, it is well-known that models which provide sufficient representation of data and simultaneously there are not too complex (have few free parameters) give good generalization. Thus, in the training process we incorporated the concept of regularization (technique for setting constraints to the weights).

Radial basis function network (RBF)

Another major class of neural network methods is the radial basis function network [4]. The RBFs are feedforward networks with one hidden layer of units. The basic characteristic of the RBF is that a hidden activation function is determined by the distance between the input vector and a prototype vector. These activation functions are called radial basis functions and give local representation of data. The most widely used basis function is the Gaussian function. As for the output units of RBFs these are typically linear.

The typical way for adjusting the weights of a RBF network for a classification problem is based on a two-stage procedure. At the first stage unsupervised techniques, such as t he k-means algorithm, are used to determine the weights associated with the basis function parameters (means and variances if Gaussian functions are used). This is achieved by considering all the training input data ignoring their class labels. At the second stage the radial basis function parameters are kept fixed and the second layer weights are adjusted using supervised learning. Typically, at this stage training is performed based on the minimization of the mean square error function with respect to the second layer weights.

In our experiments we used Gaussian functions at the hidden layer of the RBF and the first stage of training was based on probability density estimation (unsupervised technique) using Gaussian mixture models [4, 5]. The second training stage can be performed analytically (since the outputs values are linear functions of the second layer weights) by solving a linear system giving rise to mean square error solution.

The k nearest neighbor classifier (k-NN)

The k nearest neighbor classifier is a classical classification method [5]. The idea behind its use is simple: data points which are "close" under a distance measure may belong to the same class. More specifically, in order for an unknown data point to be classified, its k nearest neighbors are found from the training set (using a distance measure) and then the class with the majority over the k neighbors is chosen. In our case we used the Euclidean distance and the parameter k takes low values (3 or 5) since a few training examples are available.

Mixture models and the Bayes decision rule (GMM)

According to the statistical approach to classification [4, 5] the goal is to find the posterior probability that an unknown data point belongs to a specific class. If we denote each class as C_k $(k = 1,\dots, K)$, the posterior probability of class membership is obtained from Bayes rule

$$P(C_k \mid x) = \frac{P(C_k)p(x \mid C_k)}{\sum_i P(C_i)p(x \mid C_i)},$$

where $P(C_k)$ express the prior probability that class C_k is the true class of the data point x and $p(x \mid C_k)$ is the corresponding class conditional density function. If the prior probabilities and the class conditional densities were known, the following Bayes decision rule provides optimal results since it minimizes the misclassification error

$$\text{Choose } C_k\text{:} \quad P(C_k \mid x) > P(C_i \mid x) \quad \text{for each } i \neq k.$$

In order to use the above rule for the classification of unknown data points, first the prior probabilities and the class conditional densities have to be determined. The prior probabilities usually take equal values ($P(C_k)=1/K$) or are easily estimated from the data as $P(C_k) = N_k / N$ where N_k is the number of data belonging to class C_k and N is the total number of training data. On the other hand, the estimation of the class conditional densities using the training examples is non-trivial and techniques of probability density estimation must be applied.

In our study we model the class densities by Gaussian mixture models, which are trained using the maximum likelihood method and the EM (Expectation-Maximization) algorithm [4].

6. EXPERIMENTAL RESULTS AND CONCLUSIONS

We applied the classification methods described above separately to the data of each of the four ages. Typically we used a training set (common for all methods) to obtain the classifiers and then a test set to measure the generalization capability, that is the classification performance for unknown data points. The results are illustrated in Fig. 2 for all ages, methods and the two classification subtasks.

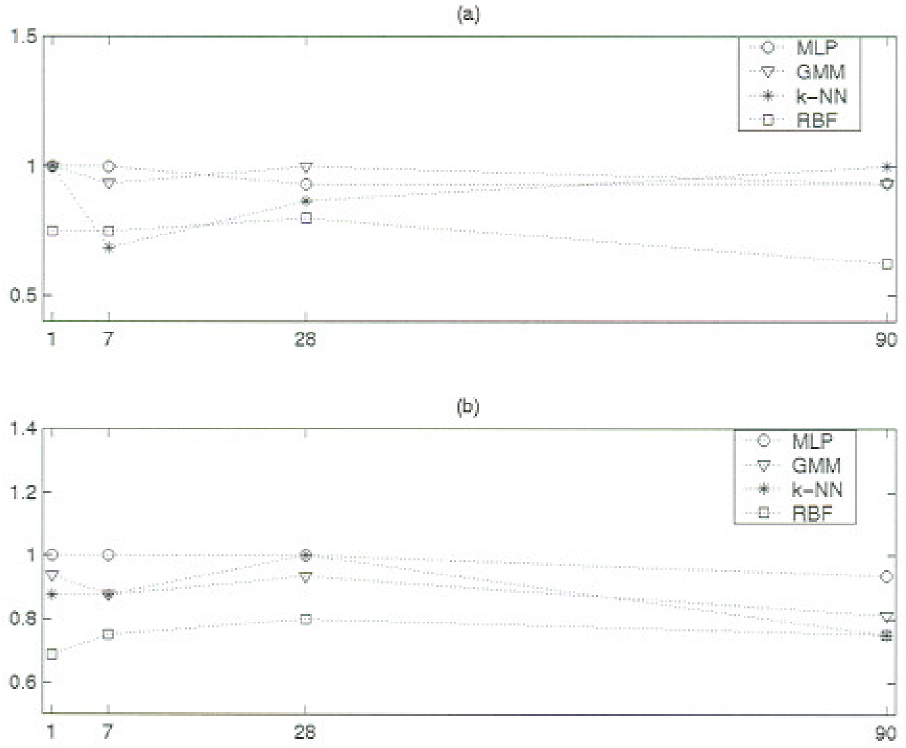

Figure 2: Representation of the generalization success rates of the methods studied for all ages: (a) results for the W/C classification problem and (b) results for the S/C classification problem.

From Fig. 2 it is obvious that some classification methods achieve high discrimination for both classification problems. The classification methods with the superior performance is the MLP and the GMM with *k*-NN follows, while the RBF had the worst performance. Remarkably, the MLP in the worst case concerning the different ages achieved success rate of 0.94 for both W/C and S/C problem.

Acknowledgements: This work is partially funded by the General Secretariat of Research and technology through the EPET II Project: "Mikkines: Non-Destructive Material Wave-analysis and Composition Monitoring".

7. REFERENCES

[1] Malhorta V. M., Carino N. J. (Eds), CRC Handbook on Nondestructive Testing of Concrete, CRC Press 1991.

[2] Vary A. The Acousto-Ultrasonic Approach, Acousto-Ultrasonics, Duke C. J. Jr. (Ed). Plenum Publishing Corporation, 1988.

[3] Αγγέλης Δ. Γ., Πολύζος Δ., Φιλιππίδης Θ. Π., Τσιμογιάννης Α., Αναστασσόπουλος Α., Γεωργαλή Β., Καλοϊδάς Β. Μη Καταστροφικός Έλεγχος Σύστασης και Εκτίμηση Αντοχής Τσιμεντοκονιάματος με την Μέθοδο Ακούστο-Υπερήχων, 1° Ελληνικό Συνέδριο Σύνθετων Υλικών Σκυροδέματος, Ξάνθη 2000.

[4] Bishop C., Neural Networks for Pattern Recognition, Oxford University Press, 1995.

[5] Duda R. O. and Hart P. E., Pattern classification and Scene Analysis, Wiley, New York, 1973.