Rough Surfaces Classification for Environmental Perception of a Mobile Robot using CTFM Sonar Imaging and Neural Networks

Konstantinos Zografos and **Penny Probert Smith**
Department of Engineering Science
University of Oxford
Parks Road, Oxford OX1 3PJ, UK
zografos,pjp@robots.ox.ac.uk

Abstract

A method for aperiodic rough surfaces classification is proposed with the use of FFT images from an ultrasound continuous transmission frequency modulated (CTFM) sensor. A mathematical model for the FFT image of the CTFM ultrasound sensor scattering from an aperiodic rough surface is derived with the help of the extended Kirchhoff Approximation Method (KAM) used by Bozma and Kuc [1]. The objective of this paper will be the discrimination of different kinds of rough surfaces, typical of a mobile robot pathway, using the attributes and the parameters of the above mathematical model for feature extraction and the help of neural networks for classification.

1 Introduction

During the past few years a number of researchers have tried to classify different kind of objects and surfaces using ultrasound sensors. Most of these efforts were based on the use of neural networks for feature extraction from the echo representation in time, frequency and time - frequency domain. In Dror et al., 1995 [2] several pre-processing transformations of the echo were examined, including power spectrum, waveform, spectrogram and cross-correlation for a three-dimensional target recognition task. The recognition of faces and the speed of a moving target using time-frequency representation were explored in Dror et al., 1996 [3]. In Harper and McKerrow [4], the classification of various plants was studied using the Fourier transform of the demodulated echo as input to a neural network. However the use of neural networks for feature extraction proves to be inefficient as the neural network manages to learn only one or two of the major features of the data. Politis and Probert [5], suggested that greater efficiency can be achieved by extracting some of the underlying features of the sonar images as pre-processing before the classifier.

In this paper the above idea of pre-processing the sonar images before the classifier, is adopted. A mathematical model for CTFM sonar scattering from rough surfaces is derived and the parameters of the model are used as features. The performance of K-Nearest-Neighbour algorithm, Radial Basis Functions and the Multi-Layer-Perceptron were tested during a classification task of different kinds of rough surfaces. Samples were taken from various kinds of rough surfaces, including grass and gravel terrain, asphalt pavement, plastic floor and a carpet, with the help of a CTFM sonar. This type of sonar transmits a linearly frequency modulated chirp and detects range in the frequency domain by comparing the instantaneous frequencies of the transmitted and received signals. The beam characteristics are similar to Polaroid sensor. Finally useful conclusions for the implementation and the robustness of each classification algorithm are extracted.

2 Rough surfaces modelling using CTFM sonar

According to Bozma and Kuc [1] the scattering of the incident wave on to a rough surface is described by the Kirchhoff approximation method. However the application of the Helmholtz-Kirchhoff integral differs from the usual treatment of approximating the incident and scattered waves by plane waves. KAM is extended to include spherically diverging waves from the scattering surface. Assuming that the surface height is described by the function $\zeta(x, y)$, which follows the Gaussian distribution with zero mean and standard deviation σ and has correlation length T, Politis and

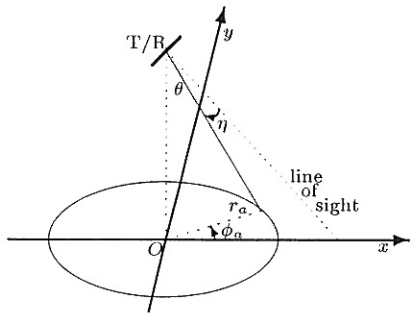

Figure 1: A reflection at a point of a circle on the surface

Probert [5] have reached a formula for direct estimation of the backscattering coefficient $\mathcal{P}$, which depends on the triplet $(\mathcal{R}, \sigma, T)$, where $\mathcal{R}$ is the absorption coefficient of the surface.

According to them when a rough surface is examined, the amplitude of the image at a specific distance will be the sum of all scattered rays of the circle on the surface plane with centre the projection point of the T/R on the surface. For all these rays, the signal attenuation in the air α and the directivity of the CTFM ultrasound sensor D are independent in terms of time and r. As it can be seen from figure 1 angle θ, which is defined as the angle between the perpendicular range and the line-of-sight, is constant on the integration line and the backscattering coefficient $\mathcal{P}(r)$ for a rough surface corresponds to all the rays emitted from T, scattered around the circle of integration and ended to R. Consequently the following equation holds for the sonar image

$$S_\alpha(r) = \frac{A^2}{4\pi\mathcal{J}}\alpha(r)\mathcal{P}(r)\int_0^{2\pi} D(\eta)d\phi_\alpha \qquad (1)$$

where $\mathcal{J}$ is the transmitter sensitivity, angle η is related to angles θ and ϕ_α and is given in [5].

So far a method has been developed for a CTFM ultrasound sensor image approximation from a rough surface. The problem of the above method is that the Thorsos criterion [6] has to be satisfied for the Kirchhoff approximation method (KAM), which already has been used in the previous algorithm, to be valid. Attempting to overcome the Thorsos criterion problem and explain the experimental results, a new statistical approach to the previous model is suggested, by splitting the rough surface into k elementary surfaces which are much smoother than the original rough surface. Each elementary surface is n_w wavelengths long, where $n_w > 1$ to avoid any conflicts with the Rayleigh

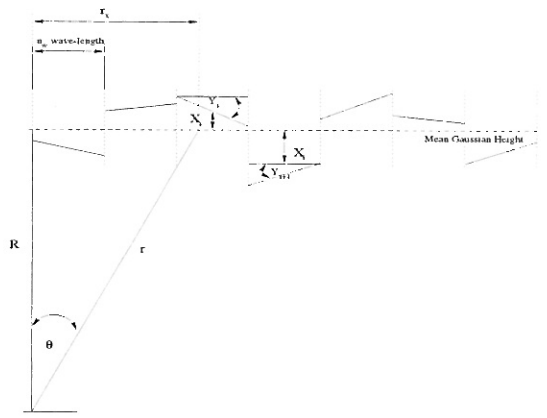

Figure 2: Modelling a rough surface using multiple reflectors

criterion, whose value decided by the simulation results to be 3, with height distributed with a Gaussian pdf of the original surface and the standard deviation depending on the roughness of the surface. The orientation of each elementary surface is random again, following a Gaussian distribution with the mean value to be 0 and the standard deviation to depend on surface roughness. The next step is to estimate the sonar image for each one of the elementary surfaces of figure 2.

At first the number n of the multiple reflectors is decided, introducing a Poisson distribution with a parameter of q. The q parameter depends on the roughness of the surface and accordingly on the fraction $\frac{\sigma}{T}$. Every reflector or elementary surface is n_w wavelengths long. If k is the number of multiple reflectors for a specific kind of rough surface and $\mathcal{X}$, $\mathcal{Y}$ are the random variables that describe the positioning and the orientation of each reflector respectively, then for each reflector denoted by i the backscattering coefficient $\mathcal{P}_i(r)$ can be estimated [5], with $r_x \in [(i-1)\cdot n_w\lambda, i\cdot n_w\lambda]$, $r = \sqrt{(R+\mathcal{X}_i)^2 + r_x}$, $\theta = \arccos\frac{R}{r} + \mathcal{Y}_i$, and i=1...k. The rough component of the sonar image is the sum of all the reflection that are coming back from all the elementary surfaces. Therefore equation 1, for the sonar image becomes

$$S_\alpha(r) = \frac{A^2}{4\pi\mathcal{J}}\alpha(r)\sum_{i=1}^{k}\mathcal{P}_i(r)\int_0^{2\pi} D(\eta)d\phi_\alpha \qquad (2)$$

where $\sum_{i=1}^{k}\mathcal{P}_i(r)$ is the summation of every $\mathcal{P}_i$, that is being produced by the i-th elementary surface. Equation 2 stands for the 'rough component' of the sonar image, while equation 1 for the 'smooth component'. A representation of the above components can be seen in figure 3. The total response from a rough surface is the combination of the two equations given by

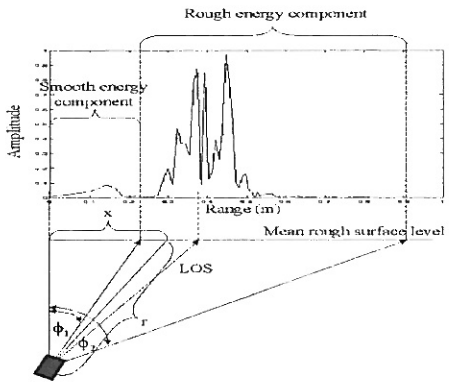

Figure 3: The 'rough' and the 'smooth' component of the sonar FFT image taken experimentally from a gravel path. The range of the x-axis corresponds to r_α of figure 1

$$S_\alpha(r) = \frac{A^2\alpha(r)}{4\pi\mathcal{J}}(\mathcal{C}_s\mathcal{P}_{smooth}(r) + \mathcal{C}_r\mathcal{P}_{rough}(r)) \int_0^{2\pi} D(\eta)d\phi_\alpha \tag{3}$$

where $\mathcal{P}_{rough}(r) = \sum_{i=1}^{k} \mathcal{P}_i$.

Theoretically the above equation is valid for every type of rough surface, because we have overcome the problem of the Thorsos criterion. Simulation results from different kinds of rough surfaces for different angles have been validated with experimental results.

3 Pattern recognition of aperiodic and periodic rough surfaces

In the previous section a mathematical model was derived. In this section the objective is practical: the discrimination of different kinds of rough surfaces based on the attributes and the parameters of the above mathematical model. As a result the focus will be on extracting the features that will provide the maximum separability among the different kinds of rough surfaces and define a class for each kind. Therefore the feature selection process will be based on the different scattering attributes that every rough surface has. Different cases of classification algorithms, like the Multi-Layer-Perceptron, the Radial Basis Functions and the K-nearest neighbour algorithm will be examined.

3.1 Feature selection for rough aperiodic surfaces

It is well established in the pattern recognition literature that a preprocessing stage of the initial data is always necessary to achieve a better performance of a classification algorithm. For the case of the CTFM ultrasound sensor scattering from rough surfaces echo preprocessing plays a crucial role in the process of classification as it turns out to be nearly impossible to train any kind of neural networks to learn the FFT and then proceed to feature extraction from the FFT image.

The task of feature extraction for the ultrasound CTFM sensor will be based on the physical evidences which can be found not only in the scattering FFT images of the CTFM sensor but within the coefficients and the attributes of the mathematical model as well.

In figure 3 the overall FFT of the scattered signal is described as the combination of two components a 'smooth' $\mathcal{P}_{smooth}$ and a 'rough' one $\mathcal{P}_{rough}$, which corresponds to the physical evidence of having a reflection component at the perpendicular direction of the scattering and one at the line of sight (figure 3). Formulating the physical and the mathematical evidence into well established features we define E_{smooth} and E_{rough} as following:

$$E_{smooth} = \int_R^{\frac{R}{\cos\phi_1}} S_\alpha(r)dr$$
$$E_{rough} = \int_{\frac{R}{\cos\phi_1}}^{\frac{R}{\cos\phi_2}} S_\alpha(r)dr \tag{4}$$

where $S_\alpha(r)$ is the FFT image amplitude at a distance r, R is the perpendicular distance from the target surface and angles ϕ_1 and ϕ_2 are dependent on the radiation pattern and the orientation of the sensor.

Another feature that is going to be proved extremely useful for the pattern recognition process is the range over which reflections are detected in the FFT image. This is related to the roughness of a surface and as it has already been discussed in section 2, it appears to be a relationship between the number of multiple targets and the roughness of a surface.

3.2 Classification of different kinds of rough surfaces

A small number of rough surfaces have been chosen for discrimination including a medium height grass terrain, a gravel, an asphalt pavement, a plastic floor and a thin corridor carpet. The selection of the above rough surfaces has been made on the basis of using the most usual surfaces that someone can find under his feet walking around the city of Oxford not only outdoors, but indoors as well. Accordingly it is very

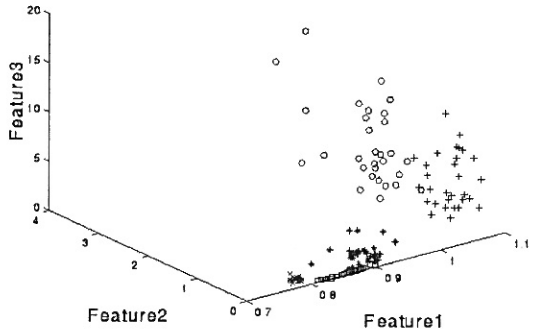

Figure 4: Class separation of gravel(cross), grass(circle), asphalt(asterisk), plastic(x-mark) and carpet(square) floor using a three dimension set of features

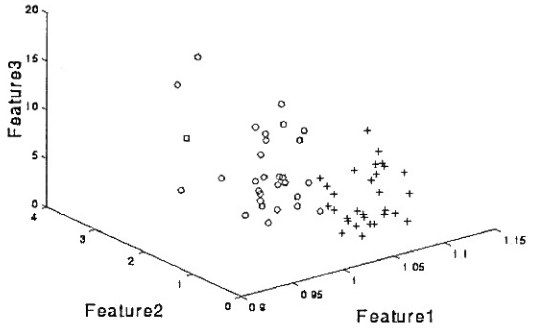

Figure 5: Class separation of gravel(crosses) and grass(circles) using a different perspective.

common to find plastic floor and carpets inside buildings, gravel and grass terrain crossing one of the many parks of the city and asphalt pavements if you walk by the side of the road.

Thirty samples were taken for each kind of the above rough surfaces. The distance and the angle from the rough surfaces were constant at 77cm and 30 degrees correspondingly. Consequently the variable values of formulae 4 can be defined as follows, $R = 0.77m$, $\phi_1 \approx 17\,\text{deg}$ and $\phi_2 \approx 50\,\text{deg}$. Although the definition of R is precise and straightforward angles ϕ_1 and ϕ_2 (see figure 3)are determined by a heuristic approximation. Apart from the features E_{smooth} and E_{rough} the range excursion over which reflections are detected in the FFT image is going to be used as well. The use of the range excursion enhances the class separation and therefore increases the robustness of the classification algorithms. The separated classes are depicted in figures 4-6.

In figure 4 a three dimensional feature space has been used to separate the classes, where feature 1 is the range over which reflections are detected in the FFT image, feature 2 is the E_{smooth} and feature 3 is the E_{rough}. As it can be observed figure 4 shows three

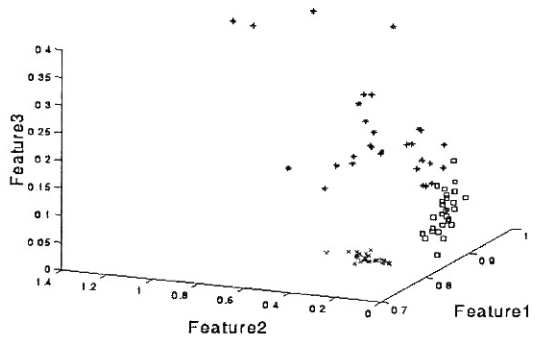

Figure 6: Class separation of asphalt(asterisks), plastic(x-marks) and carpet floor(squares) using a different perspective.

separable groups: gravel (crosses), grass (circles) and a third (x-marks, asterisks, squares), consisting of asphalt, plastic and carpet floor. If a different perspective is chosen and the third group is magnified then the classes within that group can be separated too (figure 6).

For every class a training set consisting of twenty points and a test set consisting of ten points has been chosen. Most of the training points have been chosen to be close at the boundaries of the classes to enhance the performance of all the classification algorithms that are going to be used and the generalisation process.

3.3 KNN algorithm, RBF and MLP for rough surface classification

One of the most well known and easy to implement classification algorithms is the K nearest neighbour algorithm. The number of neighbours K, acts as a smoothing parameter of the algorithm and it has a value for which the performance of the algorithm is optimal. One disadvantage of the above algorithm is that all the training data points must be retained. This might led to problems of computer storage and can require large amounts of processing to evaluate the density for new input values.

For the case of rough surface classification twenty training and ten test data points were used for each of the five classes. The results of the classification algorithm are depicted in figure 7, where the percentage of correct classification is plotted against the number K. As it can be seen from figure 7 the probability of correct classifications is significantly high for the values of K between 1 and 12 but it declines rapidly for values greater than 13.

Another well established algorithm in pattern recognition and classification uses the radial basis functions.

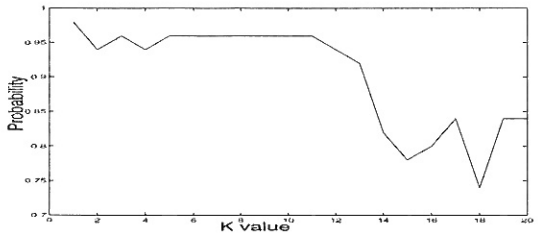

Figure 7: The correct classification probability against various values of parameter K

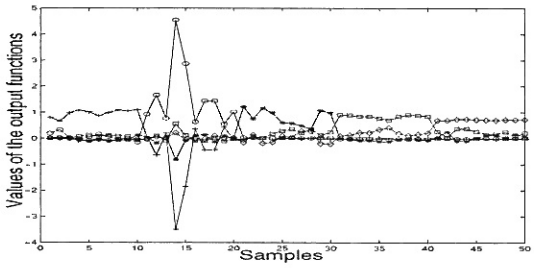

Figure 8: Output functions of the trained RBF network for the test set. The function that corresponds to the gravel test data points is denoted with crosses, for the grass with circles, for the asphalt with asterisks, for the plastic floor with squares and for the carpet with diamonds

For the above case of rough surface discrimination we choose to train a RBF network consisting of 9 hidden nodes or centres using an activation function that has the form $r^2 \log(r)$ and is known as the thin plate spline function, while the output layer is defined to be linear. The training process of the above RBF network follows two steps. During the first step the position of the centres of each radial basis function is decided with the use of the EM (expectation - maximisation) algorithm, while during the second step the hidden to output weights are determined using the pseudo-inverse method. The RBF network is trained to learn five output functions each of which corresponds to one of the five classes and takes the value 1, when data points that belong to the corresponding class are present at the input of the RBF network and 0 when they do not.

The results of the RBF training process are depicted in figure 8. The test set consists of ten data points for each one of the five classes, making a total of fifty data points. The first ten data points correspond to the first class (gravel), the second ten to the second class (grass) and so on. Therefore it is expected the first output function denoted by crosses in figure 8 to take the value 1 for the first 10 points and 0 elsewhere, the second output function denoted by circles to take the value 1 for the second 10 points and 0 elsewhere, etc. Although the output functions fail to follow the exact training values, they manage to take the greater value when a data point belongs to the class that they correspond. Consequently a 100% correct classification rate can be acquired, if a post-processing function that classifies each data point according to the function that takes the greater value is employed.

Finally the Multi Layer Perceptron is examined. A two layer perceptron was used with three inputs corresponding to the three dimensional feature space and five outputs, corresponding to the five different classes of the rough surfaces. The number of hidden nodes is four and they are activated by the logistic sigmoid function [7]. The output values of the above network are to be interpreted as probabilities, therefore they must lie in the range (0,1), and they must sum to unity. This can be achieved by using a generalisation of the logistic sigmoid activation function as the activation function for the output layer, also known as the normalised exponential, or soft-max activation function [7]. The term soft-max is used because this activation function represents a smooth version of the winner-takes-all activation model in which the unit with the largest input has output +1 while all other units have output 0.

The MLP network is trained using the scaled conjugate gradient [8] algorithm which was proved to be more robust and less computationally intensive compared with the simple optimisation method of gradient descent. The above MLP network is trained in 6780 steps using repetitively, a twenty data point train set for each of the five classes which are depicted in figures 4-6. The target values for the five output functions is 1 when a data point belongs to the class of the corresponding function and 0 when it does not, as in case of the RBF network. The overall error of the estimated output while the MLP network is being trained, converges to zero. Finally the MLP network is tested using the 50 data points test set, which has already been employed for the RBF network testing. Again, the first output function is expected to be 1 for the first ten points and 0 elsewhere, the second output function to be 1 for the second ten points and 0 elsewhere etc. As we can see in figure 9 the correct classification rate is 100 %.

4 Summary and Conclusions

One disadvantage of the K-nearest-neighbour technique is that the resulting estimate is not a true probability density since its integral over all feature space

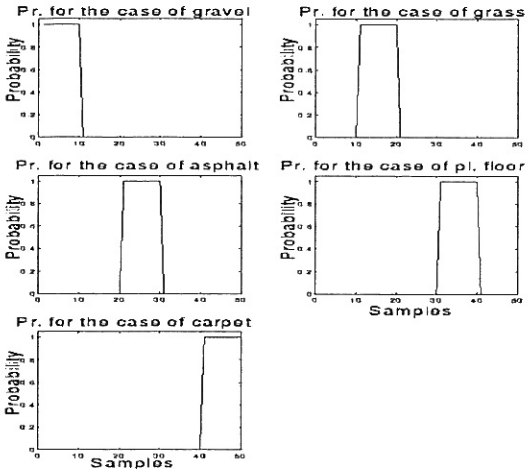

Figure 9: The five output probability functions corresponding to each one of the five classes for the rough surfaces

diverges. Another disadvantage of the above method is that all the training data points must be retained. This might lead to problems of computer storage and can require large amounts of processing to evaluate the density for new values of the input vector. Although this is not the case for the problem of rough surface classification as we have only 20 training data points for each one of the five classes, the computer storage and the processing time increases considerably as the number of classes and data points increases. For the case of rough surface classification the KNN algorithm fails to reach the 100% rate of correct classification although for a number of K values has a rate above 95%.

RBF and MLP networks play very similar role in that they both provide techniques for approximating arbitrary non-linear functional mappings between multidimensional spaces. However the particular structures of the two networks are very different. They have advantages and disadvantages which depend not only on their internal differences, but on the nature of the classification problem that they are applied to, as well.

For the task of rough surface classification the MLP network provides better results and better estimation of the output functions. The disadvantage was that it is computationally intensive as it needed 6780 steps to be trained compared with the 15 steps for the RBF network. In addition during the generalisation process the RBF network has a considerable error trying to estimate the output functions especially for the fourteenth point of the test set. That is because the RBF network has a linear output layer which degrades its performance for the given output functions and because the RBF network has defined a hyperellipsoid to describe the class of grass and the fourteenth point happens to be outside the boundaries. Apart from the fact of poor estimation of the output functions the RBF network manage to have a 100% correct classification rate if each data point of the test set is classified, based on the largest output value between the five output functions.

Finally the conclusion is that the MLP network for the rough surface classification is the best choice because of the nature of the classification problem that has to do with the definition of hyper-plane boundaries between the classes, and because of the inherent capability that the second non-linear layer provides to MLP, which is the probability estimation of any input vector to belong to a specific class. On the other hand if the computationally intensive methods are costly and if the probability estimation of an input vector to belong to a specific class is out of interest then the RBF network can be selected as it provides a very good classification rate.

Acknowledgments

One author, K. Zografos is supported by Hellenic Air Force (HAF) and Onassis Institution

References

[1] Bozma O. and Kuk R, "Characterising pulses reflected from rough surface using ultrasound," *Journal of Acoustical Society of America*, Vol. 89(6), pp. 2519-2531, 1990.

[2] Dror E. I. Zagaeski M. Moss C. F. "Three-Dimensional Target Recognition via Sonar: A Neural Network Model," *Neural Networks*, Vol. 8, pp.149-160, 1995.

[3] Dror I. Florer F. L. Rios D. Zagaeski M. "Using artificial bat sonar neural networks for complex pattern recognition: Recognizes faces and the speed of a moving target," *Biological Cybernetics*, Vol. 74, pp.331-338, 1996.

[4] Harper N. L. and McKerrow P. J. "Classification of plant species from CTFM ultrasonic range data using a neural network," *Int. Con. on Neural Networks*, pp.2346-2352, 1995.

[5] Politis Z. and Probert P. "Modelling and classification of rough surfaces using CTFM sonar imaging," *Proc. of IEEE Int. Con. on Robotics and Automation* pp. 2988-2993, 1999.

[6] Thorsos E. I. "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *Journal of Acoustical Society of America* Vol. 83, pp.78-92, 1988.

[7] Bishop M. C. "Neural Networks for Pattern Recognition," Oxford Univ. Press,1995.

[8] Moller M. "A scale conjugate gradient algorithm for fast supervised learning," *Neural Networks* Vol. 6(4), pp.525-533, 1989.