In Proceedings of the 2nd International Symposium on Autonomous Minirobots for Research and Edutainment (AMiRE), pages 219-228, Brisbane, Australia, 18 - 20 February 2003

Neural Object Classification Using Ultrasonic Spectrum Analysis

Jürgen Klahold, Hartmut Jürgens, and Ulrich Rückert

System and Circuit Technology, Heinz Nixdorf Institute,
Paderborn University, Fürstenallee 11, D-33102 Paderborn

E-mail: klahold@hni.uni-paderborn.de

Abstract. This paper describes how reflected broadband sound signals are marked by interference phenomena if the surface of the ensonified object is structured. For an efficient extraction of features of the signal that relate to the surface structure of the object, the calculation of the cepstrum is introduced. A cylindrical test object is presented, which shows an angle independent and an angle dependent structure. This allows to specify the accuracy of the discrimination of structure sizes, that is based on a selected part of the cepstrum. In addition, the object can be used as a landmark for ultrasonic sensing. Classification of the cepstra is done by the statistical 'One Nearest Neighbour' (1NN) method as well as by a 'Kohonen Self-Organising Feature Map' (SOM). The results show, that changes in structure size down to 0.1 mm are detectable.

1 Introduction

Autonomous systems are supposed to perform complex tasks under changing conditions without human intervention. Therefore, they have to interact with their environment so that the recognition of obstacles in their surrounding is essential for autonomous robots. Applying ultrasonic sensors seems to be a promising way to perceive the environment. On one hand time of flight measurements allow to estimate the position of an object [4], on the other hand the object should be identified. In this work we describe an approach of an object identification based on a ultrasonic spectrum analysis. Nevertheless it should be mentioned that an object identification is also possible in the time domain.

Ecemis [2], Fritsch [3], Abreu [7] and Lindstedt [5] use the envelope of the absolute value of the reflected signal in the time domain to gain features of the object. For feature extraction, correlation functions are used. As Lindstedt describes in his work, this method to extract features in the time domain comes along with substantial problems. Interferences may result in small amplitudes of the reflected pulse and temperature changes cause rough alternations to the received signals. Correlation functions applied to signal responses in the time domain can lead to fatal errors [5]. Good results based on the envelope of the absolute value of the reflected signal can only be achieved within these articles if specific conditions for the measurements are given.

Beside correlation functions, feature extraction is done with the classical methods of signal analysis like calculating the mean value, the maximum amplitude or the mass center of the signal. It is not stated, that these methods are chosen as the result of knowledge about relations between features of the object and features of the signal. If this knowledge would be applied, a specific signal analysis would be possible and features in the signal can be used to extract quantitative information of the object [8].

If signal analysis is based on the spectrum of a broadband signal, interference phenomena can be employed for feature extraction (sec. 2.1). Temperature changes result within the spectrum in a small shift of the interference extrema and therefore cause small errors but do not deteriorate the feature of the signal as in the time domain [5]. Streilein [9] uses 40 equidistant entries of the power spectral density of the reflected signal for classification purposes. No specific samples are chosen that are related to features of the object. Unfortunately, the features of the test object are not described in detail, hence the good classification success can not be evaluated.

McKerrow [8] uses Continous Transmission Frequency Modulated Sonar (CTFM) to evaluate distances to reflecting surfaces. The magnitude of the reflections give further information about the size of the surface. The relations between features in the signal and features of the object are used effectively. Even complex objects like plants can be classified to a certain degree with this method. For small depths of a few millimeters and small increments in depth, the CTFM reaches its boundaries since the slew rate of the sweep signal is restricted by the frequency range of the transducers and the distance to objects on one hand and by the limited resolution of the Fourier transform on the other hand.

At small structures, the interference phenomena cause rough variations in the spectrum of a reflected broadband signal (sec. 2.1). In this work it is shown, that with the help of the cepstrum an effective feature extraction of the spectrum is possible (sec. 2.2). The test object, that consists of three cylinders, provides an angle dependent and an angle independent structure. The continuous change in structure size as a function of the relative angle is given by a mathematical model. This allows to evaluate the classification successes based on spectrum analysis with the cepstrum. The angle independent structure serves as an identification feature of the object that may be used as a landmark for mobile robots.

As classification methods statistical classifiers as well as neural classifiers are applied in the described articles. The statistical classification is based on correlation functions to evaluate similarities [5] or on extracted features of the signals [8] or both [3, 7]. Feature vectors that are fed to neural classifiers are described as equidistant samples of the signal [2, 9] or are the result of prior feature extraction [8]. A comparison between both classification methods is not possible since the test objects and conditions are very different. McKerrow, who applied both methods, makes no statements about a comparison of both methods. In this work, the SOM as a representative for neural classifier and the 1NN method as a statistical classifier are applied to the acquired feature vectors. This provides the possibility to compare both classifiers.

In this work the information content of the spectrum within 20 to 120 kHz is focused. For this reason broadband electrostatic transducers are applied. Since the speed for signal processing is not the issue in this work, signal conversion is done by a digital storage oscilloscope containing a spectrum analyser that is adapted to a PC. Signal analysis, based on the output of the spectrum analyser is done with MATLAB.

2 Fundamentals

In this section we want to give a short introduction into interferences and the cepstrum, due to the fact that they are substantial for this work.

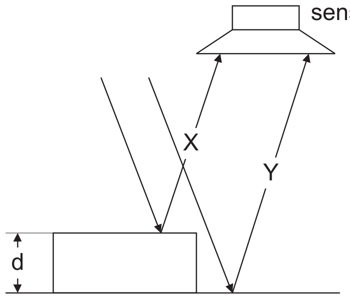

Figure 1. Reflection at a two layered structure.

2.1 Interference Phenomena

In case of a reflecting structure with two or more layers, interference phenomena occur. If there is a source that emits a sinoidal sound signal, the phase of the sound waves X and Y, as seen in (fig. 1) is the same. At a structure with two layers the wave X would suffer reflection at the forhold plane while wave Y travels further and reflects at the rear plane. When Y arrives at the forehold plane, the waves have a difference in phase according to the thickness of the structure. At the surface of the sensor both soundwaves interfere. The relationship between the frequencies of destructive interference ($f(n)$) and the structure size is given by equation 1.

$$f(n) = \frac{c}{4d}(2n+1) \tag{1}$$

Another indication for the depth of a step structure 'd' is the difference between two neighbouring interference minima (eqn. 2):

$$\Delta_f = f(n+1) - f(n) = \frac{c}{4d}(2(n+1)+1) - \frac{c}{4d}(2n+1) = \frac{c}{2d} \tag{2}$$

The amplitude of the spectrum shows a maximum if the difference in phase Δ_φ is zero or $\Delta_\varphi = 2n\pi$ and it shows a minimum if the difference in phase is $\Delta_\varphi = 2(n+1)\pi$.

The interference phenomena causes strong variations in the spectrum of the reflected sound signal. This results in problems if ultrasound systems are applied as range finders, based on one frequency. These systems send out short intervals of sound and measure the time until the reflected signal returns. If the ensonified surface contains structures that cause an interference at the used frequency, the range finder does not work [5]. On the other hand interference phenomena can be very helpful to discriminate objects. If a broadband signal is send out, like white light in optics, the reflected sound spectrum of a structured surface has a timbre, like reflections of light of a coloured surface. A method to use this effect to discriminate objects is focused in this work.

2.2 Cepstrum for Feature Extraction

The aim of this project is to find a robust and reliable method to classify objects with the help of ultrasound. In this section, a novel approach to determine the size of a step structure is discussed.

An established application for ultrasound sensors is the determination of distances to objects [7]. In the plus echo mode it is possible to determine the size of a step structure by measuring

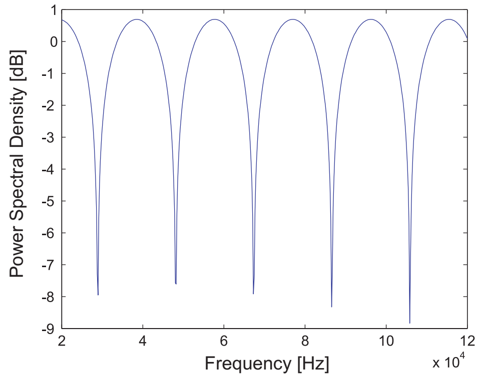

Figure 2. Spectrum of 9 mm step structure (simulation).

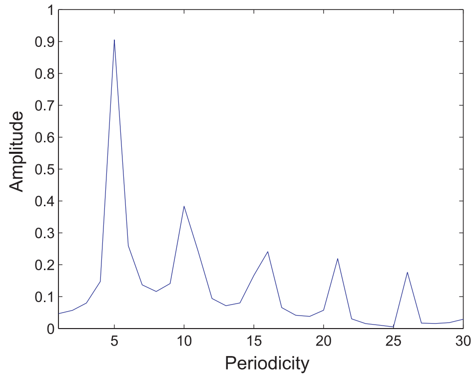

Figure 3. Cepstrum of 9 mm step structure (simulation).

the time difference between two received pulses. This however is restricted by the length of the pulses itself, hence by the bandwidth of the transducers. Another indication for the depth of a step structure is the frequency of the first destructive interference which is given by (eqn. 1) for $n = 0$.

$$f_0 = \frac{c}{4d} \tag{3}$$

The depth of a step structure d is also determined by the difference between two neighbouring interference minima or maxima as show in (eqn. 2). The spectrum of a two layered structure in dependency of the phase shift can be calculated. With the relationship between the phase angle φ and the frequency f with constant structure size (eqn. 4) the intensity I of the reflected sound results.

$$\varphi = 2\pi\frac{2d}{c}f = 4\pi\frac{d}{\lambda} \tag{4}$$

$$I = \frac{1}{Z_0}\left[a^2 + b^2 + 2ab\cos(2\pi\varphi f)\right] \tag{5}$$

Here a and b are the magnitudes of the sound wave. Equation 4 shows, that the size of the structure is represented in the rate of the periodicity of the spectrum. An increasing structure results in a higher rate of periodicity. Figure 2 shows the simulated spectrum of the reflected sound waves of a 9 mm step structure in a range from 0 to 100 kHz in dB. A method to extract the information content about the size of the step structure is to make a Fourier transform of the spectrum. As the result the periodicity within the frequency range is achieved. The highest peak in the so called ‘cepstrum’ corresponds to the depth of the step structure (fig. 3). The term cepstrum was introduced by Bogart et al. [1] and has come to be the accepted terminology for the inverse Fourier transform of the logarithm of the power spectrum ($P_{(\omega)}$) of a signal (eqn. 6).

$$C = \mathcal{F}^{-1}\left[\log(P_{(\omega)})\right] \tag{6}$$

Bogart observed that the logarithm of the power spectrum of a signal containing an echo has an additive periodic component due to the echo, and thus the Fourier transform of the logarithm of the power spectrum should exhibit a peak at the echo delay. They called this function the cepstrum, interchanging letters in the word spectrum. The cepstrum contains information about structure sizes that result in one cycle per frequency range to $n/2$ cycles, where n is the number of samples that are obtained in the frequency domain. If the structures under observation do not extend a certain depth, only the first $m \ll n/2$ entries in the cepstrum contain valuable information. This provides the possibility to reduce the data in an efficient way if only a certain range of the cepstrum is used for classification purposes.

Two effects that take place while measuring, deteriorate the cepstrum directly. These are changes in temperature and relative velocities. As mentioned in equation 2 the equation for the structure size contains the velocity of sound c as a factor, which depends on the temperature T. Equation 7 gives an approximation for sound velocity in air in a range of $-20°$C and $40°$C.

$$c_{(T)} = (331 + 0.6T)\,{}^{\mathrm{m}}\!/\!_{\mathrm{s}} \tag{7}$$

If an ultrasonic sensor is fixed to a robot, the doppler effect has to be taken into account. A relative velocity of $1\,{}^{\mathrm{m}}\!/\!_{\mathrm{s}}$ would result in a relative failure of 0.3%. A temperature change of 5 Kelvin in comparison to the reference temperature would result in a failure of 0.87%. In comparison to classification methods where correlation functions are used in the time domain, classification with the help of the cepstrum is very robust against the Doppler shift and changes in temperature. Also the problem, that interferences corrupt measurements with the ultrasound sensor systems as it is with mono frequency signals, does not occur.

The Fourier transform of the logarithm of the spectrum shows peaks beside the highest one that corresponds to the structure size. This can be overcome if the spectrum is used directly. Within this work the spectrum is calculated by a spectrum analyzer which has an output in logarithmic scale only. Experiments show, that still good results in object classification can be achieved.

3 Experiments

The aim of the project is to find alternations to the spectrum that are caused by properties of ensonified objects. The basis for successful measurements is a breadboard construction of sufficient quality. The sensors need to be sensitive to a broadband spectrum. The power supply, the transmitter and especially the receiver circuit need to guarantee low noise to achieve a high signal to noise ratio. The digital storage oscilloscope (DSO) as the link to the computer has to have a resolution that fulfills the demand of signal analysis on the personal computer. The signal generator should produce a broadband signal with a constant amplitude across all frequencies.

Ultrasonic sensors that work in air are not built to operate over a wide frequency range. The main application for these devices is to send a pulse of a single frequency, therefore the sensors operate at its resonance frequency. An acceptable frequency response over a wide frequency range provides the Polaroid sensor of the 600 series.

Autonomous robots, like the minirobot Khepera [6], need to determine their position to fulfill tasks. While moving the robot can calculate its current position by counter that adds up the spins of the wheels. The limited accuracy of this procedure result in massive failure [6]. To recalibrate the system, information about the distance and the angle to a known location is necessary. This task can be preformed by a landmark which provides an angle and distant dependent reflection (fig. 4). As seen in figure 5 the size of a structure can be determined with a high accuracy if the cepstrum (fig. 6) of a reflected broadband spectrum is used for feature extraction. An object with an angle depending structure should therefore allow to conclude from the reflected spectrum to the relative angle to the landmark. The distance to the landmark can be determined by measuring the time of flight of the sent signal. To discriminate the landmark against other objects, a feature for the identification is necessary. This can be provided by an angle independent structure. Result of this investigation is an object containing an angel

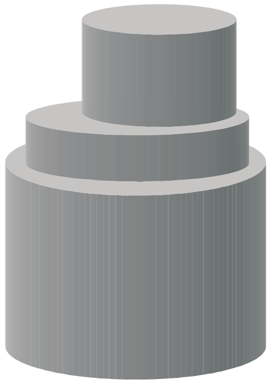

Figure 4. Cylindrical Landmark.

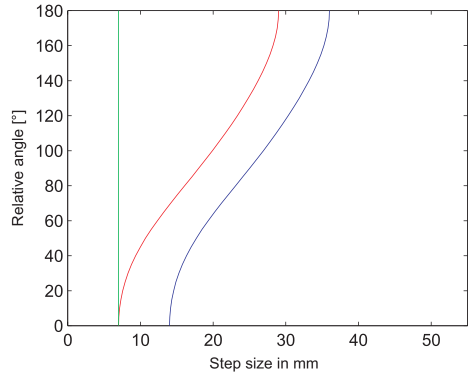

Figure 5. Structures of the landmark.

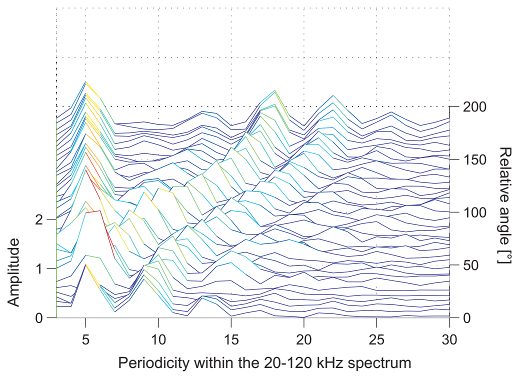

Figure 6. Cepstra of the landmark at 30 cm distance.

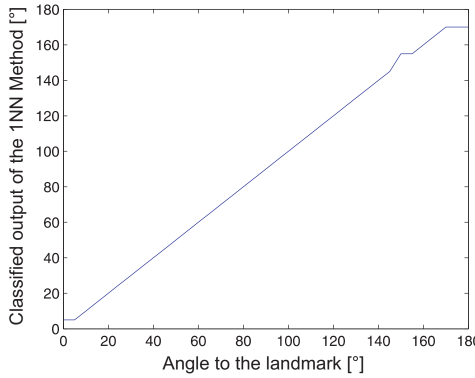

Figure 7. 1NN classification: 30 cm(1) against 30 cm(2).

depending and an angle independent structure, which is given by three cylinders, arranged in a manner as shown in (fig. 4).

A reference about the dimension of the object that reflects sound energy is given by the fact, that objects with an extent of more than five wave length produce a distinct acoustical shadow. For this reason the diameter of the smallest cylinder d_i is chosen according to the lowest frequency of the broadband signal of 20 kHz to 85 mm.

The identification structure is chosen to 7 mm, which results in a periodicity of 4 within a frequency range of 100 kHz. This peak is found sufficiently far from peaks that are caused by the offset the characteristic of the sensor and the Hamming window function. The angle dependent structure is chosen to 7 mm as the smallest- and to 29 mm as the biggest size. This guarantees a high gradient in structure size which is the premise to discriminate cepstra of neighbouring angles.

4 Classification of Measured Data

In this section the classification methods 1NN and SOM are applied to the measured data. Classification is done with simulated data of the mathematical model and with measured data of two different data-sets from three distances (30 cm, 60 cm and 90 cm).

The 1NN method applied to the two first sets of measurements at 30 cm distance (fig. 7) shows only one misclassification of 5° at 160° within 5° and 170°. At the margins of the scale the

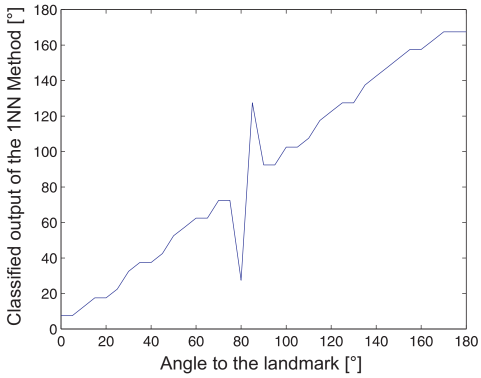

Figure 8. 1NN classification: 30 cm(1) against 30 cm(3).

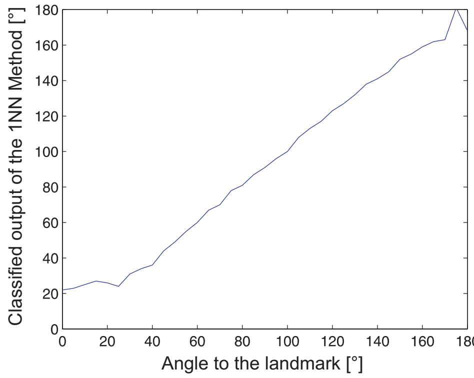

Figure 9. 1NN classification: 60 cm (simulated) against 90 cm(2).

classification error does not exceed 10°. Misclassification at the margins results of the low gradient in change in structure size at these angles of the landmark.

Since the measurements are taken at two different places under different conditions, the discrimination between the angle of 5° and 10° can be considered as the highest resolution that is achieved with this breadboard construction if the 1NN method is applied (fig. 7). The corresponding change in structure size can be calculated and results to 0.1027 mm. The best resolution in terms of angle occurs at 90°. Here the structure changes within 85° and 95° by 1.9243 mm. The best resolution in this area would be 0.53°.

The third data-set of 30 cm distance contains cepstra of 2.5°, 7.5°, ... while all other data-sets contain the cepstra of 5°, 10°, ... For classification this is the worst case situation. Classification of the first two data-sets of 30 cm distance with the third measurements of this distance results in two rough misclassifications (fig. 8). Beside that, the errors within 5° and 170° do not exceed a range of 2.5° in both cases.

If two data-sets of one distance are applied to the 1NN method, the results are generally good. Failures occur at the edges at 0° and 180° respectively where the angle dependent changes are small. Slightly worse are the results with combinations of data-sets of 60 cm and 90 cm (fig. 9). Combinations where the data-set of 30 cm distance is used with data-sets of other distances lead to unsatisfying outputs.

With a SOM of 7×7 elements, data can be classified without rough misclassification. A faultless classification would result in entries that are located on a straight line with a gradient of one which passes through the origin. Those identifiers that are displayed as a star represent the classification results of the training-data. Those that are displayed as a circle refer to the classification results of the test-data. While test-data is presented it happens, that neurons are winners that are not active while training-data is used. These neurons are unlabelled and lead to an error-identifier of '-10'.

Figures 10 and 11 shows typical classification results of the SOM. In this case the SOM is trained with a data-set of 37 cepstra, measured at a distance of 30 cm and 60 cm in room 1 respectively. The test-data that is classified consists of 37 cepstra as well, measured at a distance of 30 cm and 90 cm in room 2 respectively. At the edges, the cepstra within about 20° are often assigned to one class, since changes in the cepstra in this range are small. In this case a misclassification at 20° leads to the result, that the test-data is assigned to a labelled neighbouring neuron. This causes a failure within the magnitude of the resolution of the SOM and can be tolerated. Generally, good classification is possible if the training-data and the test-data of the

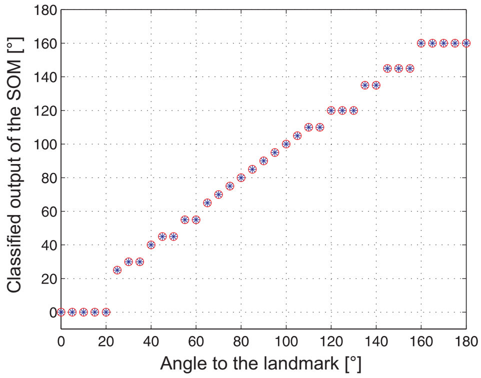

Figure 10. SOM classification: 30 cm(1) training-data, 30 cm(2) test-data.

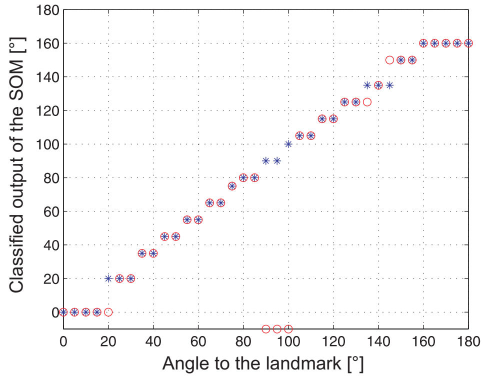

Figure 11. SOM classification: 60 cm(1) training-data, 90 cm(2) test-data.

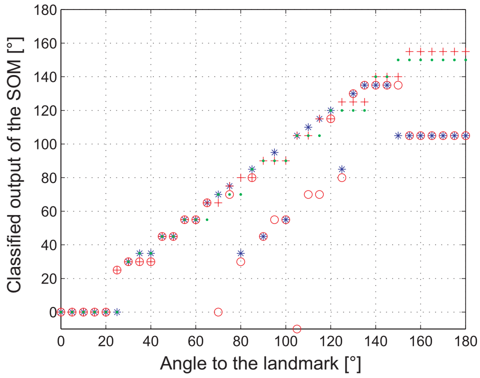

Figure 12. SOM classification with training-data of room 2 of 30 cm, 60 cm, 90 cm distance and test-data of room 1 of 30 cm distance.

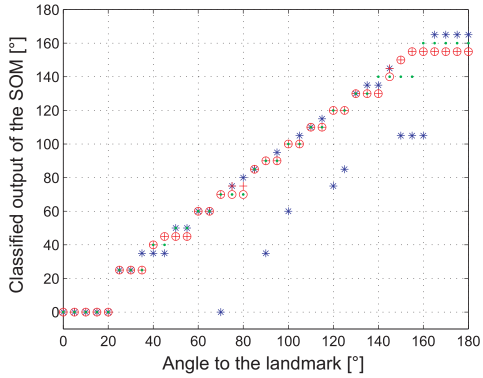

Figure 13. SOM classification with training-data of room 2 at 30 cm, 60 cm, 90 cm distance and test-data of room 1 of 60 cm distance.

same distance are taken. Acceptable results occur if data of 60 cm distance is classified with data of 90 cm distance and vice versa.

An option that potentially leads to good results for all distances lies in training the SOM with data of 30 cm, 60 cm and 90 cm distance. It can be expected that less error-identifier are found, since more neurons are labelled. The classification success should be worse, since a part of the training-data differs roughly to the test-data. The output of the test-data of 30 cm (fig. 12) and of 60 cm (fig. 13) are displayed as red circles. In both cases the neurons represent more feature vectors which results in a rough classification. In case of training-data of 30 cm distance big errors often occur. The errors at the upper margin is caused by the association of the active neuron to the smallest angle of its class, which in this case is 105°. More serious are the offset errors that can only be detected to a certain degree by a system that is based on continuous measurements. Compared to classification results based on 60 cm or 90 cm training-data the use of training-data of all distances gives an output with less errors. The result for the 60 cm test-data shows rarely rough misclassifications or errors that result of unlabelled neurons.

All together the output of the SOM that is trained with the feature vectors of all distances leads to better results if feature vectors of small and big distances are classified. To obtain a classification which is optimised over all regarded distances, it is also possible to use the time of flight to evaluate the distance to the landmark. Afterwards a distance optimised classifier could

range where error$\leq 5°$ (error in mm)	30 cm room 2	60 cm room 2	90 cm room 2
30 cm, room 1	$0° - 175°$ (0.14 mm)	-	-
60 cm, room 1	-	$15° - 175°$ (0.26 mm)	$15° - 175°$ (0.26 mm)
60 cm, room 2	-	0	$35° - 155°$ (0.58 mm)

Table 1. Classification accuracy of the 1NN method.

range where error$\leq 10°$ (error in mm)	30 cm room 2	60 cm room 2	90 cm room 2
30 cm, room 1	$25° - 170°$ (0.98 mm)	-	-
60 cm, room 1	-	$25° - 170°$ (0.98 mm)	-
60 cm, room 2	-	0	$35° - 155°$ (1.23 mm)

Table 2. Classification accuracy of the SOM.

be chosen.

5 Results

A comparison of the SOM and the 1NN method for this application shows, that the statistical classifier shows better results than the neural classifier. The reason for that is, that the classifiers of the 1NN method are given by the feature vectors itself. The SOM has to be trained so that the classifiers move to the position of the feature vectors that defines a class. A comparison of the 1NN method and the SOM is possible with the tables 1 and 2. Fields with the entry '-' express that these measurements have too many errors. The entry '0' expresses that this measurement would make no sense since classifiers and test-data are based on one set of data. The entries with the unit 'mm' stands for the accuracy in structure size at the margin with the bigger gradient. All entries of the table with the results of the 1NN method are at least as twice as accurate as the results of the SOM.

6 Conclusion

The Fourier transform of the logarithmic spectrum, the cepstrum proved to be an effective tool for feature extraction (sec. 2.2). Peaks in the cepstrum can be related directly to the size of surface structure of the object. The zoning of an array in the cepstrum according to the size of the relevant surface structures enables an efficient compression of information.

Classification of the reflected ultrasound signals of the landmark is realised by the 'One Nearest Neighbour Method' as a representative of statistical classifiers and by a two dimensional 'Self-Organising Feature Map' as a representative of neural classifiers. During tests with 37 feature vectors it turns out, that good results occur if the SOM is chosen to a quadratic grid of 49 neurons. Less neurons result in bigger classes and more neurons in more errors during the test phase.

Best results are gained with the 1NN method at 30 cm distance (fig. 7). The minimum change in structure size that is detectable at this distance is 0.1027 mm. With increasing distance the accuracy decreases, but the robustness against misclassification of feature vectors of other distances increases. Robust classification with only few errors and high accuracy are measured at 60 cm distance. A statistic of 10 measurements lead to an average accuracy of 4.95° with a standard deviation of 6.07° within 0° and 180° and of 2.52° with a standard deviation of 4.05° within 20° and 160°. Twice misclassification occurred which led to an error-identifier. These data-sets are not used in this statistic.

A comparison of both classification methods shows, that the 1NN method produces much better results than the SOM (tab. 1 and tab. 2). The use of a SOM for feature extraction makes sense if unsupervised learning is required.

The size of the structures determine the periodicity within a certain spectrum, which again determines the accuracy of the system. This points out the limitations of the method. A periodicity of one within a spectrum of 100 kHz corresponds to a structure size of 1.7 mm. Especially for small structures an alternative method in spectrum analysis should be used. A possibility is to find its first or second or third . . . interference minimum. These frequencies can also be used to determine the structure size.

The gained information form the cepstrum could be used for an ultrasonic based object classification. Furthermore, the accurate detection of structure sizes is also useful for a global positioning system as mentioned in this article.

References

[1] **B. P. Bogart**, **M. J. R. Healy**, and **W. Tukey**. The Quefrency Analysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Saphe Cracking. In *Proc. of the Symposium on Time Series Analysis*, pp. 209–243, Brown University, Providence, USA, 1962.

[2] **M. Ihsan Ecemis** and **P. Gaudiano**. Object recogntion with ultrasonic sensors. In *Proc. of the IEEE Int. Symposium on Computational Intelligence on Robotics and Automation (CIRA)*, pp. 250–255, 1999.

[3] **C. Fritsch**, **J. J. Anaya**, **A. Ruiz**, and **L. G. Ullate**. A high-resolution object recognition ultrasonic system. *Sensors and Actuators A*, 37–38:644–650, 1993.

[4] **J. Klahold**, **J. Rautenberg**, and **U. Rückert**. Continuous Sonar Sensing for Mobile Mini-Robots. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA02)*, pp. 323–328, 2002.

[5] **G. Lindstedt**. Recognition using Frequency Sweep Sonar. In *IEEE Proc. of the 1st Euromicro Workshop on Advanced Mobile Robots (EUROBOT)*, pp. 46–53, 1996.

[6] **A. Löffler**, **J. Klahold**, and **U. Rückert**. The Mini-Robot Khepera as a Foraging Animate: Synthesis and Analysis of Behaviour. In *Proc. of the 5th Int. Heinz Nixdorf Symposium: Autonomous Minirobots for Research and Edutainment (AMiRE)*, volume 97 of *HNI-Verlagsschriftenreihe*, pp. 93–130, 2001.

[7] **J. M. Martín Abreu**, **T. Freire Bastos**, and **L. Calderón**. Ultrasonic echoes from complex surfaces: an application to object recognition. *Sensors and Actuators A*, 31:182–187, 1992.

[8] **J. P. McKerrow** and **L. N. Harper**. Recognising leafly plants with in-air sonar. *Sensor Review*, 19(3):202–206, 1999.

[9] **W. W. Streilein**, **P. Gaudiano**, and **G. A. Carpenter**. A Neural Network for Object Recognition through Sonar on a Mobile Robot. In *Proc. of the IEEE Int. Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, pp. 271–276, 1998.