Object Classification and Localization with an Airborne Ultrasound Imaging System

Wei Yap Tan[1], Grischan Erbacher[1], Till Steiner[2], and Nicole V. Ruiter[1]

[1] *Karlsruhe Institute of Technology, Germany*
Email: wei.tan@kit.edu
[2] *Pepperl+Fuchs GmbH, Germany*

Abstract

An airborne ultrasound imaging system (ABUS) was developed at KIT for reflection tomography. The prototype system consists of sixteen ultrasonic transducers surrounding a region of interest (ROI) of defined shape with a diameter of 50 cm. The transducers have a center frequency of 200 kHz and a bandwidth of 20 kHz. The prototype aims to demonstrate possible industrial applications for object classification and localization with airborne ultrasound. This paper is a detailed version of the previous publication in IEEE Ultrasonics Symposium (IUS) 2017 [1].

Keywords: airborne ultrasound, reflection tomography, object classification, object localization

1 Introduction

An airborne ultrasound imaging system (ABUS) was developed at KIT for reflection tomography. The prototype system consists of sixteen ultrasonic transducers surrounding a region of interest (ROI) of defined shape with a diameter of 50 cm. The transducers have a center frequency of 200 kHz and a bandwidth of 20 kHz. The transducer distribution is optimized for a maximum coverage of the custom ROI and an homogeneous distribution of image quality using the optimization method described in [2]. In the current setup, each transducer sends an ultrasound pulse and the received signals, i.e. A-scans, at all transducers are sampled. In contrast to recent publications using linear ultrasound arrays [3, 4], this system is capable of imaging multiple objects from multiple angles around the ROI in a single acquisition using synthetic aperture focusing technique (SAFT)[5]. The reconstructed images allow recognition and localization of multiple objects simultaneously. Figure 1 shows the current system setup.

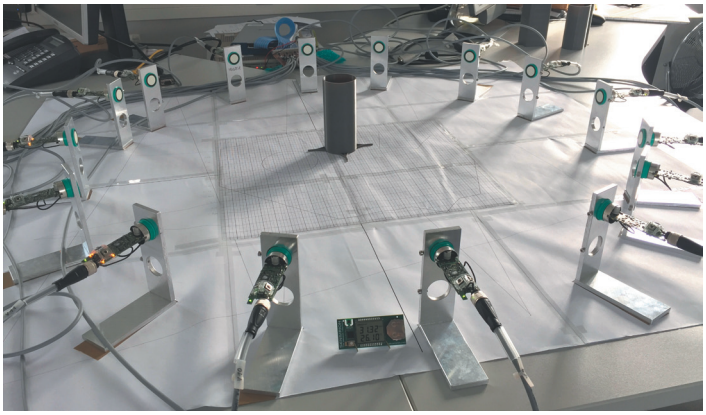


Figure 1: ABUS with sixteen ultrasonic transducers surrounding the ROI

In this work we aimed to automatically segment, classify and localize objects in the reconstructed image. The resulting workflow is essential in demonstrating possible applications for such airborne ultrasound imaging systems. For instance, ABUS can assist the localization of transparent objects, or objects of different colors with the same shape, where a camera solution faces difficulties. Figure 2 shows an example application of ABUS in industry, in which the robot arm can pick the desired objects based on the results from the proposed workflow.

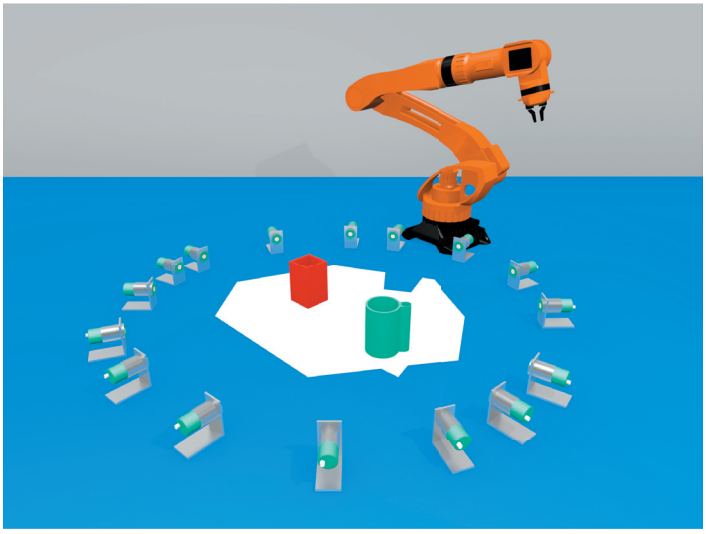


Figure 2: The figure shows a potential industrial application of ABUS for object recognition and localization.

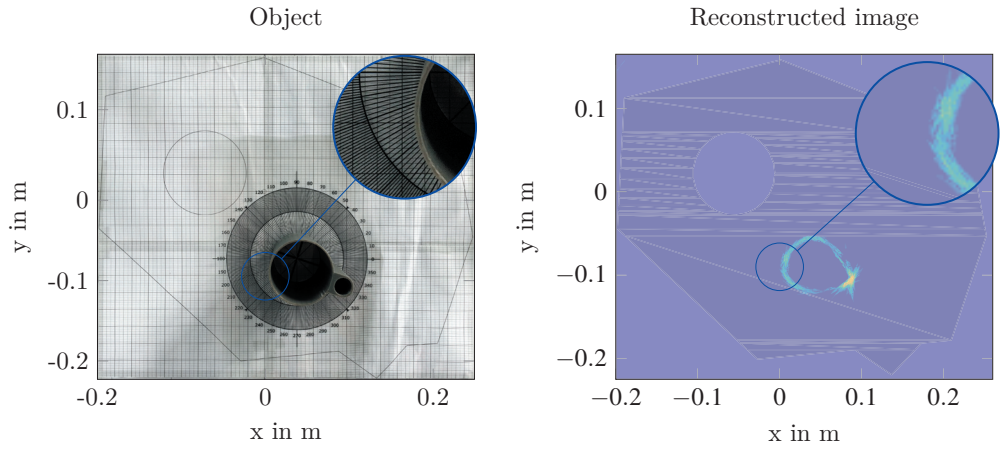


Figure 3: The images show the comparison between a camera image (left) and a reconstructed ultrasound image with ABUS (right).

2 Challenges and Objectives

Figure 3 shows an example reconstruction of an object with an asymmetrical feature. In the reconstructed image the object surfaces are depicted as edges with varying intensities. Besides that, the reconstructed edges contain spiky artifacts as shown in the magnified image. Due to the sparse transducer arrangement and the low bandwidth of our current setup, objects positioned closely to each other might not always be separable in the reconstructed images.

At the same time, large acoustic impedance between air and solid objects limits the transmission of the ultrasound wave through the objects. This leads to acoustic shadowing that results in incomplete reconstruction of objects despite multiple acquisition angles around the ROI. These image characteristics restrict the direct application of state of the art pixel oriented, edge and cluster based segmentation techniques, such as thresholding, active contours and region-growing methods.

Our objective is to develop an automatic workflow, which can be applied for the varying intensities of object edges to segment, classify and localize the objects. Our goal is to achieve a classification accuracy higher than 90 % and a localization accuracy comparable to the full width at half maximum (FWHM) of the point spread function (PSF) of 1.5 cm of the example system.

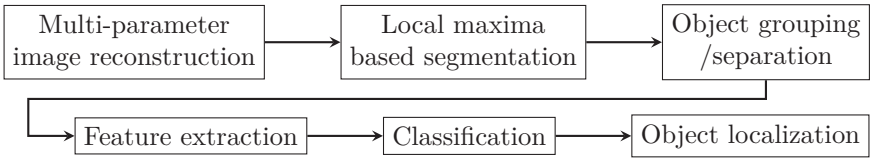


Figure 4: Workflow for automatic object segmentation and classification

3 Methods

Figure 4 shows the proposed workflow for automatic segmentation and object classification with ABUS. The workflow is a sequential process consisting of image reconstruction, image segmentation, post-processing of the segmentation results and classification and localization of the segmented objects.

3.1 Multi-Parameter Image Reconstruction

Before the SAFT reconstruction, we applied echo detection to the sampled A-scans and replaced the detected echoes with a sinc pulse with an equivalent FWHM of 1.7 mm, which is the wavelength of the ultrasound frequency used. Using SAFT, the intensity of each pixel is a sum of the signal amplitude in the A-scans for each selected sender-receiver-combination at the time equal to the time-of-flight for an ultrasound wave traveling from the sender to the pixel and to the receiver. In order to reduce grating-lobe artifacts, we extended our SAFT reconstruction algorithm with the sliding median filter introduced in [6].

According to this reconstruction method, the reconstructed image details and contrast can be influenced by changing the sets of sender and receivers and the width of the sliding median filter used. The two parameters are the angle ϕ, which limits the amount of receivers per sender and the median width k. The optimal values of these parameters are dependent on the size and surface curvature of the objects and their position and orientation in the ROI.

For example, a cylindrical object with a small radius reflects the ultrasound wave over a larger angle compared to a plane surface. The echoes will be received by more receivers nearby the sender, hence larger ϕ and k can be used. This settings are less suitable for reflection at plane surface, as the reflection will be treated as outliers by the median filter. Figure 5 demonstrates schematically the SAFT reconstruction with sliding median in ABUS.

The influence of ϕ and k on the image details is shown in Figure 6 with the same test object as in Figure 3. With a larger angle $\phi = 80^\circ$ and median width $k = 5$, the spiky artifacts of the reconstructed object were reduced. However these parameters introduced an incomplete edge, as compared to the reconstruction with $\phi = 40^\circ$ and $k = 3$. Since we have full access to the

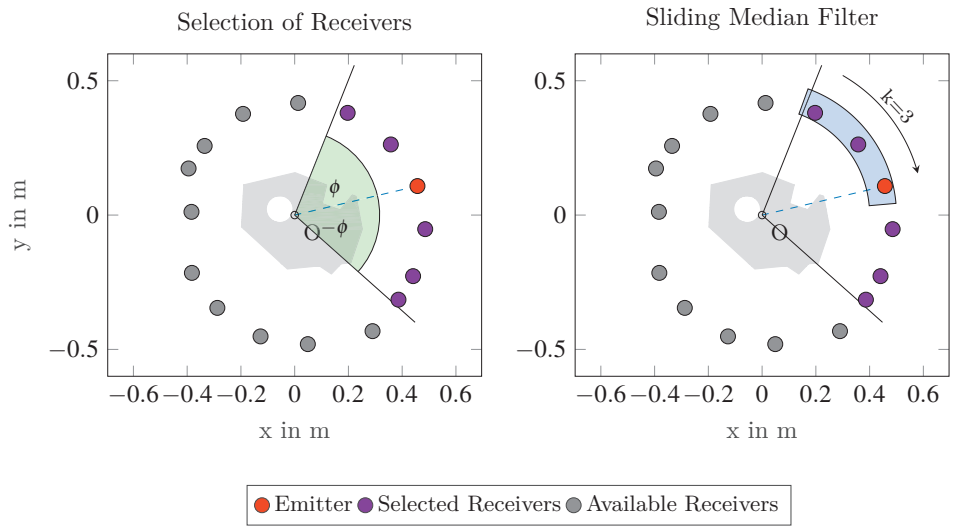


Figure 5: The figures show the selection of receivers for the Emitter S1 during the SAFT reconstruction with angle ϕ and the application of a sliding median for grating lobe reduction with a width of $k = 3$.

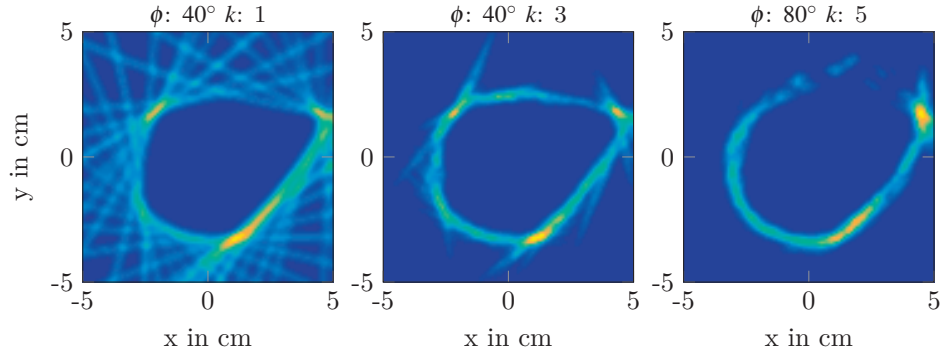


Figure 6: Influence of angle ϕ and median filter length k on the image details

A-scans, we propose the reconstruction of multiple images using different parameters to obtain better segmentation for objects with different sizes and surface curvatures.

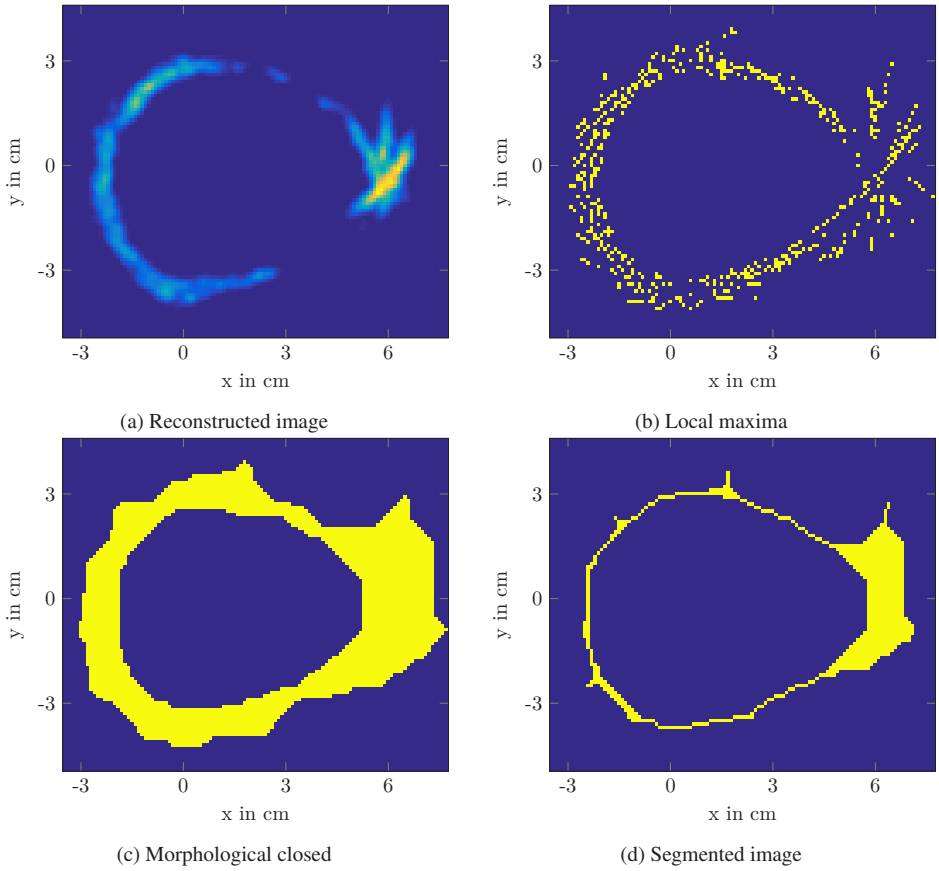
3.2 Local Maxima based Segmentation

The varying intensity of the reconstructed objects limits the application of a global threshold to segment the image. Hence, the object segmentation applies local maximum detection using

a sliding Gaussian window. The width of the Gaussian window can be set according to the measured point spread function (PSF) of the ultrasound imaging system.

In a separate matrix of the same size as the image, the frequency of a pixel being detected as a local maximum is recorded. Local maxima with occurrence frequencies lower than 3 were empirically found to be image artifacts and are discarded. Next, the remaining local maxima with intensities higher than the 98th percentile are selected, resulting in the image in Figure 6(b). It is generally assumed that higher intensity pixels are more likely to be part of the object edges and the object edges cover only a small amount of pixels in the images.

Local maxima extracted from the reconstructed images with different parameters are combined and morphological closing is applied to turn these local maxima into connected regions as shown in Figure 6(c). Lastly, the morphological skeletons of these regions are calculated to obtain the segmented object in Figure 6(d).



(a) Reconstructed image

(b) Local maxima

(c) Morphological closed

(d) Segmented image

Figure 7: The figure shows the steps of an example segmentation.

3.3 Object Grouping or Separation

Due to the sparse transducer arrangement and shadowing effect of ultrasound at obstacles, an object may be reconstructed incompletely and segmented as separate regions. Post-processing of the segmentation results can be performed using prior knowledge about the imaging capability of the ultrasound system to join these regions to a single object.

In this work, we applied the Euclidean distance between pixels of each segmented region to group regions belonging to the same object. The maximum distance, in which two regions are grouped as a single object, can be set according to the PSF of the system or the minimum size of objects to be imaged.

3.4 Feature Extraction and Classification

The classification of the segmented object is implemented with a neural network for pattern recognition [7]. In this work, the following features are extracted from the segmented object as input values for the classification:

- Area
- Convex area
- Major axis length
- Minor axis length
- Eccentricity
- Perimeter
- Roundness metric
- Squareness metric
- Within clusters sum of point-to-centroid distances from k-means clustering [8].

The roundness metric R is calculated with

$$R = \frac{4 \cdot \pi \cdot \text{Area}}{\text{Perimeter}^2}. \tag{1}$$

Whereas the squareness metric S is calculated as

$$S = \frac{\text{Area}}{16 \cdot \text{Perimeter}^2}. \tag{2}$$

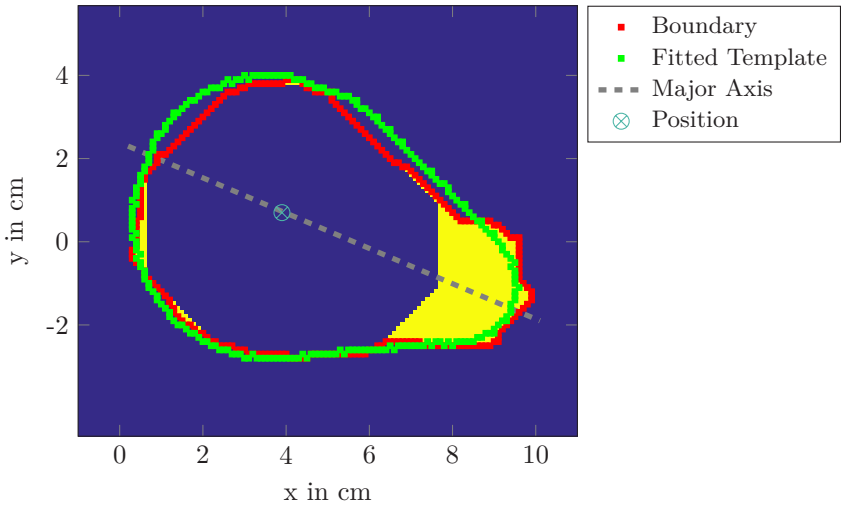


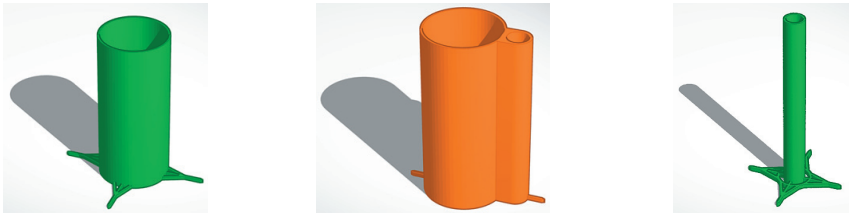
Figure 8: Object localization with Generalized Hough Transform

For a set of objects to be trained, we extract these features from reconstructions containing only one object and train the neural network under supervised learning with Bayesian Regularization Backpropagation, which prevents overfitting of the network [9].

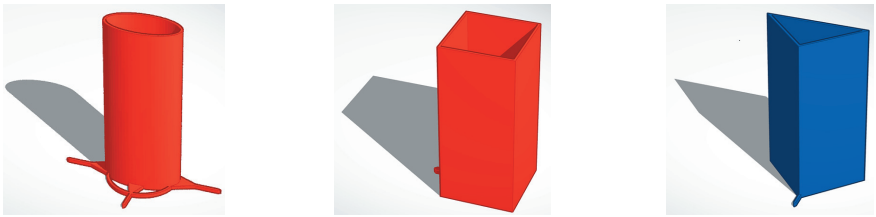
3.5 Object Localization

The position and orientation of a segmented object can be determined by calculating the centroid and the major axis of the segmented region. For more complex objects with known geometry, template matching techniques such as Generalized Hough Transform (GHT) can be applied [10]. In our work, each object in the data set is assigned with a customized template, which is a binary image consisting of the outline of the object.

Figure 8 shows an example of the object localization with a template. The yellow area shows the segmented region with its outer boundary in red. The template is rotated during the application of GHT, in order to enable the detection of the object in different orientations. The position and angle of the template with the largest maximum point in the Hough space indicate the location and orientation of the object. In Figure 8, the fitted template is plotted with a green line.



(a) Cylinder 1 (b) Asymmetrical cylinder (c) Cylinder 2



(d) Elliptic cylinder (e) Cuboid (f) Triangular prism

Figure 9: Test objects for evaluating the automatic segmentation, classification and localization of objects with ABUS

4 Evaluation

For the evaluation of the proposed workflow, six objects of different shapes and sizes were designed. These objects have in general a diameter of 6 cm and a height of 12 cm except for the smaller cylinder with a diameter of 1.5 cm (Cylinder 2 in Figure 9). The cross section of the objects are primitive shapes, such as circles, ellipses, rectangles, triangles and circles with an asymmetrical feature as shown in Figure 9. The shown CAD models were used for 3D printing of the test objects.

The multi-parameter image reconstruction was performed with angles $\phi = \{60°, 80°, 100°\}$ and $k = \{5, 6, 7\}$. For each object, forty measurements were performed at discrete positions and angles in the ROI. This data was divided into 70 % for the training of the neural network and 15 % each for validation and testing. The neural network has a hidden layer with fifteen neurons and an output layer with six neurons. The resulting overall confusion matrix after eighteen iterations is shown in Figure 10. A classification accuracy of 98.8 % was achieved in the training phase, 86.1 % in the validation phase and 88.9 % in the test. The overall classification accuracy was 95.4 %.

The asymmetrical cylinder was used as the test object in evaluating the localization accuracy. A total of sixty acquisitions was performed at discrete position and orientation and the ground truth values were recorded. We used the template as shown in Figure 8 for the GHT and obtained a positioning accuracy of 5 mm with a standard deviation of 2.7 mm and orientation accuracy of 2.8° with a standard deviation of 5.4°.

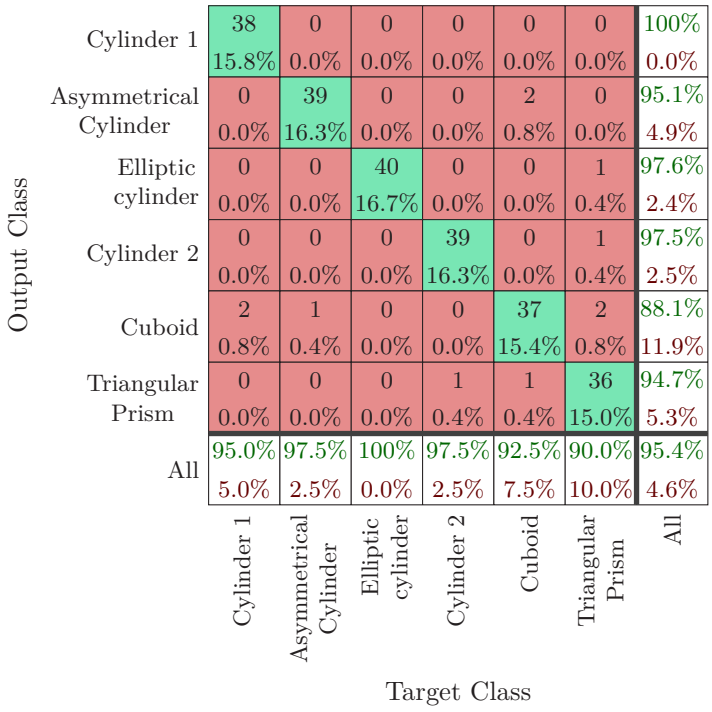


Figure 10: Confusion matrix of the trained neural network for classifying the six test objects

5 Conclusion

In this work, we introduced an exemplary workflow for automatic object segmentation and classification for an airborne ultrasound imaging system. The local maxima based segmentation provides good results despite varying intensity and PSF in the reconstructed images. In combination with the multi-parameter SAFT, we were able to detect and classify objects of different sizes and surface curvature. Figure 11 shows the ability of the workflow in classifying and localizing two objects simultaneously with ABUS.

The evaluation results showed a overall classification accuracy higher than the required 90 % at 95.4 % with an accuracy of 88.9 % in the test phase. A localization accuracy of 5 mm was achieved, which is smaller than the measured FHWM of PSF of 1.5 cm. These results fulfilled the given objectives for our workflow.

Due to the sparse transducer arrangement and the limited bandwidth of the current prototype, the segmentation results cannot differentiate object features smaller than approximately 2 cm. This could be further improved with usage of more transducers or higher bandwidth of ultrasound transducers.

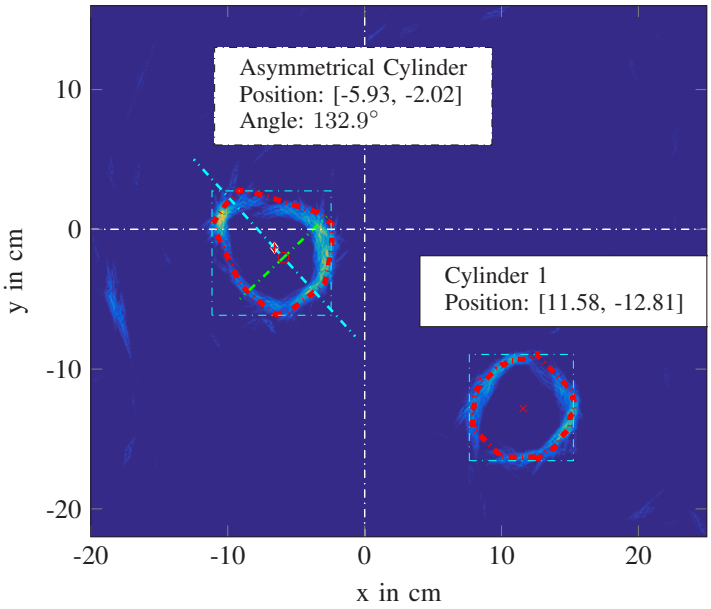


Figure 11: The figure demonstrates the ability of the workflow to classify and localize two objects simultaneously with ABUS.

References

[1] Tan, W. Y., Erbacher, G., Steiner, T., and Ruiter, N. V. Automatic segmentation and object classification with neural network for an airborne ultrasound imaging system. In *Ultrasonics Symposium (IUS), 2017 IEEE International* (2017).

[2] Tan, W. Y., Steiner, T., and Ruiter, N. V. Automatic optimization of sensor positioning for an airborne ultrasound imaging system. In *Ultrasonics Symposium (IUS), 2016 IEEE International* (10 2016).

[3] Brown, G., and Reilly, D. Proceedings of ultrasonics international 1995 ultrasonic tomographic imaging of solid objects in air using an array of fan-shaped-beam electrostatic transducers. *Ultrasonics 34*, 2 (1996), 111 – 115.

[4] Capineri, L., Bulletti, A., Calzolai, M., and Giannelli, P. An airborne ultrasonic imaging system based on 16 elements: 150 khz piezopolymer transducer arrays—preliminary simulated and experimental results for cylindrical targets detection. *Sensing and Imaging 17*, 1 (2016), 1–23.

[5] Stepinski, T. An implementation of synthetic aperture focusing technique in frequency domain. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 54*, 7 (7 2007), 1399–1408.

[6] Ruiter, N. V., Dapp, R., Zapf, M., and Gemmeke, H. A new method for grating lobe reduction for 3d synthetic aperture imaging with ultrasound computer tomography. In *2010 IEEE International Ultrasonics Symposium* (10 2010), pp. 2372–2375.

[7] Bishop, C. *Neural Networks for Pattern Recognition*. Advanced Texts in Econometrics. Clarendon Press, 1995.

[8] Wu, J. *Advances in K-means Clustering: A Data Mining Thinking*. Springer Publishing Company, Incorporated, 2012.

[9] Burden, F., and Winkler, D. *Bayesian Regularization of Neural Networks*. Humana Press, 2009, pp. 23–42.

[10] Ballard, D. H. Readings in computer vision: Issues, problems, principles, and paradigms. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987, ch. Generalizing the Hough Transform to Detect Arbitrary Shapes, pp. 714–725.