Francesca Tonetti, BSc

Development of a software tool to analyse the accuracy of algorithms implemented in acoustic cameras

Master's Thesis

to achieve the university degree of

Master of Science

Master's degree programme: Elektrotechnik-Toningenieur

submitted to

Graz University of Technology

Supervisor

Ph.D. Luca Fredianelli, IPCF-CNR, Pisa
Gaetano Licitra, ARPAT, Pisa
Ao.Univ.-Prof.i.R. Dipl.-Ing. Dr.techn. Gerhard Graber, SPSC TU Graz
Dipl.-Ing. Dr.techn. Werner Weselak, SPSC TU Graz

Institute for Signal Processing and Speech Communication
Head: Univ.-Prof. Dipl-Ing. Dr.techn. Gernot Kubin

Graz, September 2023

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Date	Signature

Abstract

For the past few decades, acoustic cameras have been utilized to locate noise sources and visualize them through the use of beamforming techniques. As interest in this field has grown, scientists and companies have developed new antenna and microphone array designs and algorithms to enhance the accuracy and efficiency of this method. However, the benefits of each technique remain unclear both theoretically and practically. This thesis examines the results of applying various beamforming algorithms to one acoustic camera measurements in controlled conditions. The study revealed a fundamental implementation error in a commercially available acoustic camera and the still arising imprecisions at low frequencies.

Kurzfassung

In den letzten Jahrzehnten wurden akustische Kameras und Beamforming-Arrays zur Geräuschquellenlokalisation und -visualisierung eingesetzt. Nach dem Interessenswachstum in diesem Bereich haben Wissenschaftler und Unternehmen neue Antennen- und Mikrofonanordnungen und Algorithmen entwickelt, um die Genauigkeit und Effizienz dieser Anwendung zu verbessern. Die Vorteile der einzelnen Techniken bleiben jedoch sowohl theoretisch als auch praktisch unklar. Diese Arbeit untersucht die Ergebnisse der Anwendung verschiedener Beamforming-Algorithmen auf eine akustische Kameramessung unter kontrollierten Bedingungen. Die Untersuchung enthüllte einen grundlegenden Implementierungsfehler in einer handelsüblichen akustischen Kamera und bestätigte die immer noch auftretenden Ungenauigkeiten bei niedrigen Frequenzen.

Bibliography

[1] *Noise definition*. [Online]. Available: https://www.lexico.com/definition/noise (cit. on p. 1).

[2] M. Bolognese, F. Fidecaro, D. Palazzuoli, and G. Licitra, "Port noise and complaints in the North Tyrrhenian sea and framework for remediation," *Environments - MDPI*, vol. 7, no. 2, pp. 1–17, 2020, ISSN: 20763298. DOI: 10.3390/environments7020017 (cit. on p. 1).

[3] M. Nastasi, L. Fredianelli, M. Bernardini, L. Teti, F. Fidecaro, and G. Licitra, "Parameters affecting noise emitted by ships moving in port areas," *Sustainability (Switzerland)*, vol. 12, no. 20, pp. 1–17, 2020, ISSN: 20711050. DOI: 10.3390/su12208742 (cit. on p. 1).

[4] *Interreg*. [Online]. Available: https://interreg-maritime.eu/fr/web/pc-marittimo/home (cit. on p. 1).

[5] *LIST*. [Online]. Available: https://interreg-maritime.eu/web/listport (cit. on p. 1).

[6] *TRIPLO*. [Online]. Available: https://interreg-maritime.eu/web/triplo (cit. on p. 1).

[7] *Rumble*. [Online]. Available: https://rumble-project.eu/i/ (cit. on p. 1).

[8] *Mon Acumen*. [Online]. Available: https://interreg-maritime.eu/web/monacumen (cit. on p. 1).

[9] *Porto di Livorno Wiki*. [Online]. Available: https://it.wikipedia.org/wiki/Porto_di_Livorno#Caratteristiche (cit. on p. 2).

[10] Autoritá di sistema portuale, "movimenti_portuali_2018," Tech. Rep., 2018 (cit. on p. 2).

[11] A. Badino, T. Gaggero, E. Rizzuto, D. Borelli, T. Gaggero, E. Rizzuto, and C. Schenone, "Acoustical Impact of the Ship Source The 21 st International Congress on Sound and Vibration ACOUSTICAL IMPACT OF THE SHIP SOURCE," (cit. on p. 4).

[12] A. Badino, D. Borelli, T. Gaggero, E. Rizzuto, and C. Schenone, "Airborne noise emissions from ships: Experimental characterization of the source and propagation over land," *Applied Acoustics*, vol. 104, pp. 158–171, Mar. 2016, ISSN: 1872910X. DOI: 10.1016/J.APACOUST.2015.11.005 (cit. on p. 4).

[13] L. Fredianelli, M. Nastasi, M. Bernardini, F. Fidecaro, and G. Licitra, "Pass-by characterization of noise emitted by different categories of seagoing ships in ports," *Sustainability (Switzerland)*, vol. 10, no. 5, pp. 1–12, 2020, ISSN: 20711050. DOI: 10.3390/su12051740 (cit. on p. 5).

[14] J. Billingsley and R. Kinns˜, "The acoustic telescope," Tech. Rep. 4, 1976, pp. 485–510 (cit. on pp. 6, 15).

[15] J. Christensen and J. Hald, "Technical Review Beamforming," Brüel & Kjær, Naerum, Tech. Rep., 2004 (cit. on pp. 6, 10).

[16] J. Grythe, "Beamforming algorithms-beamformers," Norsonic AS, Oslo, Tech. Rep., 2015 (cit. on pp. 9–12).

[17] L. Segers, B. Da Silva, A. Braeken, and A. Touhafi, "Cloud-Based Acoustic Beamforming Emulator for FPGA-Based Sound Source Localization," in *2018 4th International Conference on Cloud Computing Technologies and Applications, Cloudtech 2018*, Institute of Electrical and Electronics Engineers Inc., Jul. 2018, ISBN: 9781728116372. DOI: 10.1109/CloudTech.2018.8713349 (cit. on p. 9).

[18] G. Heilmann and D. Döbler, "Perspectives of the Acoustic Camera," Tech. Rep., 2005. [Online]. Available: https://www.researchgate.net/publication/235957582 (cit. on p. 9).

[19] Wolfram Research, "ConjugateTranspose," Wolfram Research, Champaign, Illinois, USA, Tech. Rep., 2004. [Online]. Available: https://reference.wolfram.com/language/ref/ConjugateTranspose.html (cit. on p. 12).

[20] P. Stoica, Z. Wang, S. Member, J. Li, and S. Member, "Robust Capon beamforming; Robust Capon beamforming," *IEEE Signal Processing Letters*, vol. 10, no. 6, 2003. DOI: 10.1109/LSP.2003.811637. [Online]. Available: http://www.sal.ufl.edu/wang/RCBDiagLoad.pdf (cit. on pp. 12, 13).

[21] J. A. Högbom, "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines," in *Astronomy & Astrophysics*, vol. 15, 1974, p. 417 (cit. on p. 13).

[22] P. Sijtsma, "CLEAN based on spatial source coherence," *13th AIAA/CEAS Aeroacoustics Conference (28th AIAA Aeroacoustics Conference)*, 2007, ISSN: 1475-472X. DOI: 10.1260/147547207783359459 (cit. on p. 13).

[23] U. Rau and T. J. Cornwell, "A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry," *Astronomy & Astrophysics*, vol. 532, A71, Aug. 2011, ISSN: 0004-6361. DOI: 10.1051/0004-6361/201117104 (cit. on p. 13).

[24] P. Gupta and S. P. Kar, "MUSIC and improved MUSIC algorithm to estimate direction of arrival," in *2015 International Conference on Communication and Signal Processing, ICCSP 2015*, Institute of Electrical and Electronics Engineers Inc., Nov. 2015, pp. 757–761, ISBN: 9781479980819. DOI: 10.1109/ICCSP.2015.7322593 (cit. on p. 14).

[25] R. P. Dougherty, "FUNCTIONAL BEAMFORMING," Tech. Rep. (cit. on p. 14).

[26] U. Michel, "History of acoustic beamforming," *Proceedings of the Berlin Beamforming Conference*, pp. 1–17, 2006. [Online]. Available: http://elib.dlr.de/47021/1/BeBeC_2006_Paper_Michel.pdf (cit. on p. 15).

[27] Robert P. and Dougherty, *Spiral-shaped array for broadband imaging*. [Online]. Available: https://www.freepatentsonline.com/5838284.pdf (cit. on p. 16).

[28] g. tech GmbH, "Sound Localisation with the Acoustic Camera," [Online]. Available: www.gfaitech.com (cit. on p. 17).

[29] "SORAMA CAM64 Specification sheet," [Online]. Available: www.sorama.eu (cit. on p. 17).

[30] "Sorama CAM iV64 Product sheet," [Online]. Available: www.sorama.eu (cit. on p. 17).

[31] "SORAMA CAM1K Specification sheet," [Online]. Available: www.sorama.eu (cit. on p. 17).

[32] *Sevenbel*. [Online]. Available: https://www.sevenbel.com/en/ (cit. on pp. 17–19).

[33] *Norsonic Acoustic camera*. [Online]. Available: https://web2.norsonic.com/product_single/acoustic-camera/ (cit. on pp. 17, 18).

[34] *BK Connect Acoustic Camera — Brüel & Kjær*. [Online]. Available: https://www.bksv.com/en/instruments/daq-data-acquisition/analyzer-system/acoustic-camera-9712 (cit. on pp. 17, 18).

[35] *Akustische Kamera - Schallkamera — CAE Systems*. [Online]. Available: https://www.cae-systems.de/produkte/akustische-kamera-schallquellen-lokalisieren.html (cit. on pp. 18, 20).

[36] Gfai, *fmin*. [Online]. Available: https://www.gfaitech.com/knowledge/faq/frequency-range-microphone-array (cit. on p. 26).

[37] J. Martin Bland and D. G. Altman, "STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT," *The Lancet*, vol. 327, no. 8476, pp. 307–310, Feb. 1986, ISSN: 01406736. DOI: 10.1016/S0140-6736(86)90837-8 (cit. on p. 35).

[38] *Propagation of Error*. [Online]. Available: https://it.wikipedia.org/wiki/Propagazione_degli_errori#:~:text=In%20fisica%2C%20per%20propagazione%20degli,ad%20una%20funzione%20di%20esso. (cit. on pp. 38, 42).

Contents

1 Introduction

Noise [ŋɔɪʒ]: a sound that is loud or unpleasant and causes disturbance [1].
Noise pollution and its effect on human health has been studied for some years now and researches have shown that long term exposure to high noises can cause severe damage. On a long term sight humans can encounter cardiovascular and respiratory disorders and hypertension as physical harm and concentration and hearing loss as psychical harm [2]. These symptoms are often the consequence of sleep deficiency. Most noise pollution research has been analysing traffic noise caused by cars, trains and airplanes mainly. Only in recent years, as maritime transport for commercial shipping and transportation around the world increased significantly, the focus also shifted towards ship noise and industrial ports [3]. Latter represent a complex environment with manifold working processes, which provoke a very diverse noise environment. The whole port area consists of an extensive network including trains, ships, heavy vehicles and industrial machinery. Each one produces its own noise while following the tight daily schedules to assure the efficiency of maritime commerce transport. However, only in recent times, awareness from citizens in several South European port areas like Livorno, Genova, Nice increased. At this time neither in Italy nor in France, specific decrees have been published. The European Noise Directive 2002/49/EC (END) tries to force its members towards a progressive reduction of noise with the help of action plans and up-to-date noise maps of road, rail and air traffic and industries [2]. Therefore, there is a standardized methodology and procedure for creating rail and road noise maps but once again, not enough awareness of the consisting problem in port areas is raised.
The INTERREG Maritime Programme Italy-France 2014-2020 [4] included numerous projects with the aim of understanding port noise in the north Tyrrhenian Sea for long-term sustainability. These strategies included DECIBEL for noise reduction infrastructures, L.I.S.T. PORT [5] for vehicle traffic, TRIPLO [6] for noise in port and logistic areas, RUMBLE [7] for noise emission mitigation and MON ACUMEN [8] for noise monitoring and sustainability management. The complexity and variance of noise in the port area, including ship noises which vary by type and activity as described in chapter 1.1.2 from page 4 pose a major challenge. Despite these challenges, combining measurements can

produce noise maps showing average and peak noise levels, which must remain below specified levels to assure general human well-being.
Initially the fundamental aim of this work was to identify and subsequently exactly locate the main noise source in such a conglomerate. The measurements were planned to be made with a sound level meter for static and enduring noises and the acoustic camera for moving and momentary noises in a port of the northern Tyrrhenian Sea, specifically the port of Livorno. This two approaches should provide enough data to create accurate noise maps. The organization of those measurements was the first obstacle. The port of Livorno is one of the largest ports in the Mediterranean Sea. Considering the areas within and adjacent the customs barriers used for port activities (including warehouses, storage rooms, etc.) it expands over an area of 2.500.000 m^2 [9]. Over 60 km of railway lines connect the port to the main line *Livorno - Pisa*. In 2018, freight traffic alone accounted for 37 million tonnes [10]. Numerous industries from mainly international shipping companies (ATLAS s.r.l., Lorenzini) to big ferry lines (Grimaldi Lines, Moby), to chemical industries (Solvay), to engineering companies (Laviosa) are present. To enter their sites of activity a personal permit for each visitor is required. Additionally, the exact date of entrance has to be announced and since all the measurements were conducted open air, the weather conditions were another determining factor. Also, the ship traffic, loading and unloading of the cargo ships is precisely timed within a few hours and the permits cannot be issued at short notice. Once the team managed to enter the private sites it had to select the optimal positioning for the equipment. Especially in the freight areas the setup of the apparatus was complicated as there are strict safety measures to adhere to. Only after the arrangement of the necessary equipment and a first orientation, the diverseness of the multiple sounds began to unfold. The following chapter 1.1 will draw a picture of the distinct noise sources which collectively cause the noise pollution around the port area.
As the area of interest turned out to be more complex than expected, the exact scientific use of the acoustic camera and subsequently the data analysis settings in the corresponding software for the purpose were not clear enough. For this reason the team decided to do some research about beam-forming theory and acoustic cameras development. This fundamental aspects will be discussed in section 1.2 and 1.3.
To investigate the acquired theoretical knowledge a new experiment in a more controlled and easier to access environment was designed. The whole process from the idea, over the completion of the measurements to the data analysis tool development is described in chapter 2. Chapter number 3 contains an analysis and discussion of the graphical results. The last chapter, namely

number 4 includes a summarising conclusion with a significant observation about the functionality of one of the implemented algorithms. Furthermore, the concluding chapter illustrates an idea for future scientific work.

1.1 Acoustic situation of an industrial port

One of the largest European industrial ports extends over a surface of $2.500.000m^2$. The system consists of vast parts on land with factory buildings, heavy machinery such as pipes, cranes, pumps and stocking containers along with a complex transport routes network with trains and trucks. Additionally, the water traffic must be taken into account. The following subsections 1.1.1 and 1.1.2 describe the influence of each industrial element on noise pollution.

1.1.1 Noise pollution caused by the industries

Numerous industries are present in the port of Livorno. International and national transportation companies like *Sintermar, Lorenzini&C. and Darsena Toscana* are responsible for loading and unloading cargo ships, freight trains or trucks. Fixed port cranes load each container onto a truck which drives to the scheduled deposit position. Then a reach stacker piles them up or unloads them onto the transfer train. This workflow takes several hours for one cargo ship and often ships are already waiting outside the port. Thus these industries are operating during the whole day and constantly producing different heavy machinery noise.
Another industrial category are chemical or oil industries like *Laviosa Chimica, D'Alesia Group, Eni Gas costiero* and *Industrie Levora*. These companies make use of large pumps, tractor units, trucks and also forklifts. Each of these machines adds to the noise level in the area.
Additionally, many small land transport businesses have their headquarters and contribute to the traffic noise emission considerably.
Lastly, the shipyards like *Cantiere Navale Il Calambrone, Mare Bleu* and *MEC-Carpensalda di Lotti Bruno & C.* represent a huge noise source. Building and repairing ships includes movement of fixed cranes or moving cranes; hammering, drilling, screwing and polishing of large metal pieces and transportation of heavy parts.
The noise of each of this enterprises adds up to a substantial amount of disturbance.

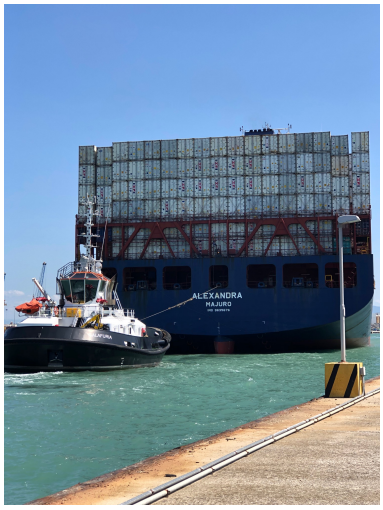


Figure 1.1: Fully loaded cargo ship

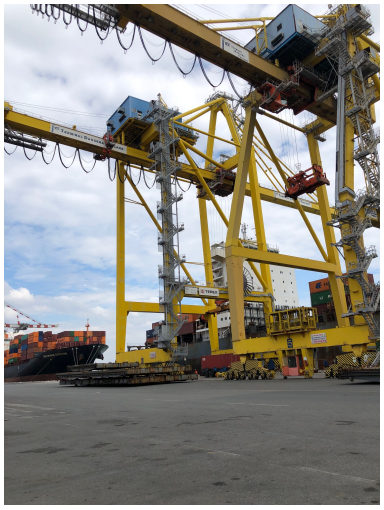


Figure 1.2: Loading of a ferry boat

Figure 1.3: Industry

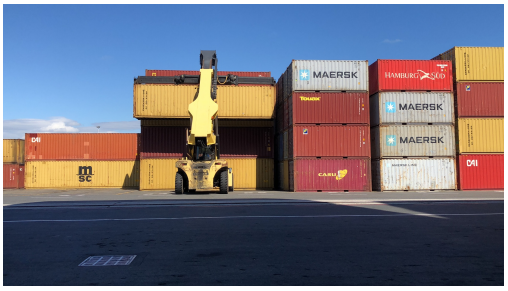


Figure 1.4: Loading of a ferry boat

1.1.2 Noise pollution caused by ships

In general ships represent a complex noise source. There are on board noises attributed to propulsion and exterior radiation through air and water related to the movement of the ship [11]. The first resembles typical moving objects like road vehicles, airplanes and trains. However, when anchored it is more comparable to anthropogenic noise sources like industrial plants. In fact, Badino et Al. [12] subdivided the ships' movements into four categories. The first are ships sailing close to the coast at a reduced speed for harbor approach or channel pass through. The second are ships maneuvering for entering (or exiting) ports, the third are moored ships without any ongoing activities and

the fourth category are ships at quay with loading or unloading processes. These distinct actions involve particular working conditions for the vessels. For maneuvers the motors work in exceptional conditions and often auxiliary equipment such as bow and stern thrusters, azimuthal propulsors and winches are involved and augment the resulting noise level. At berth the main engines are not running. However, generators for electricity, ventilators for garages (ferries) and air-conditioning (passenger boats) are in operation. Through the ships' funnel gases are dissipated at a high speed and add to the totality of the noise. Considering the numerous chimneys along the vessel the noisy area around the boat increases considerably. The last element to acknowledge is the loading and unloading of cargo or passenger ships. Cars, trucks and other vehicles on ramps and cranes are commonly involved and cause impulsive noise. To sum up, there are broadband, long-lasting noise elements combined with short, frequency-specific impulses.
For the industrial port of Livorno, Fredianelli et al. [13] divided the passing ships into five categories:

i) Roll-on/roll-off (RORO) [49%]
ii) container ships [16%]
iii) oil tanker [20%]
iv) chemical tanker [5%]
v) ferry [9%]
vi) other [1%]

Data collected with the help of a weather mast and class 1 sound level meter in combination with the surveillance cameras of the Port Authority made it possible to calculate the average sound power level (L_w), the third octave band sound power spectrum and the duration of transition for each category. Subsequently the noise impact of each vessel type could be classified. The results can be seen in table 1.1.

Table 1.1: Duration of transition in seconds and average sound power level in A-weighted decibel when entering the industrial port of Livorno.

Category of Vessel	Transition duration [s]	L_w [dB(A)]
RORO	103	87.5
Container ships	120	89
Oil tanker	118	82.6
Chemical tanker	104	85.9
Ferry	109	84.3

These noise levels can be accounted to the different ventilation systems over the

ships' hull, the engines, the ventilation around the chimneys and no to neglect the water motion caused by the vessels.

(a) Fully loaded cargo ship (b) Ferry boat transit (c) Loading of a ferry boat

Figure 1.5: Three different types of vessels in the industrial port of Livorno.

1.2 Beamforming

The use of a microphone array to individuate the origin of a noise was introduced for the first time by Billingsley and Kinns in 1976 [14]. Back then the aim was to acquire knowledge about the sound field of a turbojet engine to subsequently make it quieter. From then on, beamforming has been applied to many different fields such as high-speed trains, cars in motion, flying aircrafts and rotors like helicopters or wind turbines. The rapid development of processing capabilities has led to a vast increase in sampling frequencies. As a result, significantly larger amounts of data can be processed at higher resolutions. Overall, this enhanced the accuracy of results, and beamforming nowadays is indispensable for localising sound sources.

1.2.1 Principles of Beamforming

The following considerations aim to introduce to the complex topic of beamforming.
Beamforming is an array-based measurement technique for sound source localization from medium to long measurement distances [15]. The distance is a key-factor because the localization is done by time delay Δt between the single microphones and is only valid for the plane wave observation (see figure 1.6). The plane wave front approaches the array with an unknown angle α, depending of its direction of arrival (DOA). The distance $d_{m,min}$ between the single microphones is known and the time delay between the microphones

can be estimated. Considering the known speed of sound in air c, the time difference can also be expressed as a length l per speed c.

$$\Delta t = \frac{l}{c} \tag{1.1}$$

The length l is the distance from the second microphone to the wave front at the moment the wave reaches the first microphone and can be calculated as:

$$l = d_{m,min} \cdot \sin \alpha \tag{1.2}$$

Combining the equations 1.1 and 1.2 the direction of arrival α can be calculated.

$$\alpha = \arcsin \frac{\Delta t \cdot c}{d_{m,min}} \tag{1.3}$$

At the same time the array size, more specifically the distance between the microphones restricts the usable frequency range of the acoustic camera. For sound waves with a wavelength smaller than twice the distance between microphones $d_{m,min}$ the unequivocal localisation of the source can not be garanteed:

$$d_{m,min} \leq \frac{\lambda_{min}}{2} \tag{1.4}$$

with $d_{m,min}$ being the smallest distance between the microphones. Reshaping the formula and substituting λ with $\frac{c}{f}$ leads to:

$$f_{max} = \frac{c}{2d_{m,min}} \tag{1.5}$$

and accordingly the theoretical lower frequency limit is given with:

$$f_{min} = \frac{c}{2d_{m,max}} \tag{1.6}$$

Nonetheless, besides this fundamental beamforming principles there are several different algorithms deployed for DOA estimation. This work only focuses on the five algorithms put to work during the practical free field experiment: *Delay and Sum (DaS)*, *Capon*, *Clean SC*, *Multiple Signal Classification (MUSIC)*, *EigenValue-Optimized Beamforming (EVOB)*. Each of these algorithms can be allocated to a distinct beamforming methodology. The simplest algorithm *DaS* is based on conventional beamforming and is characterized by a fixed steering

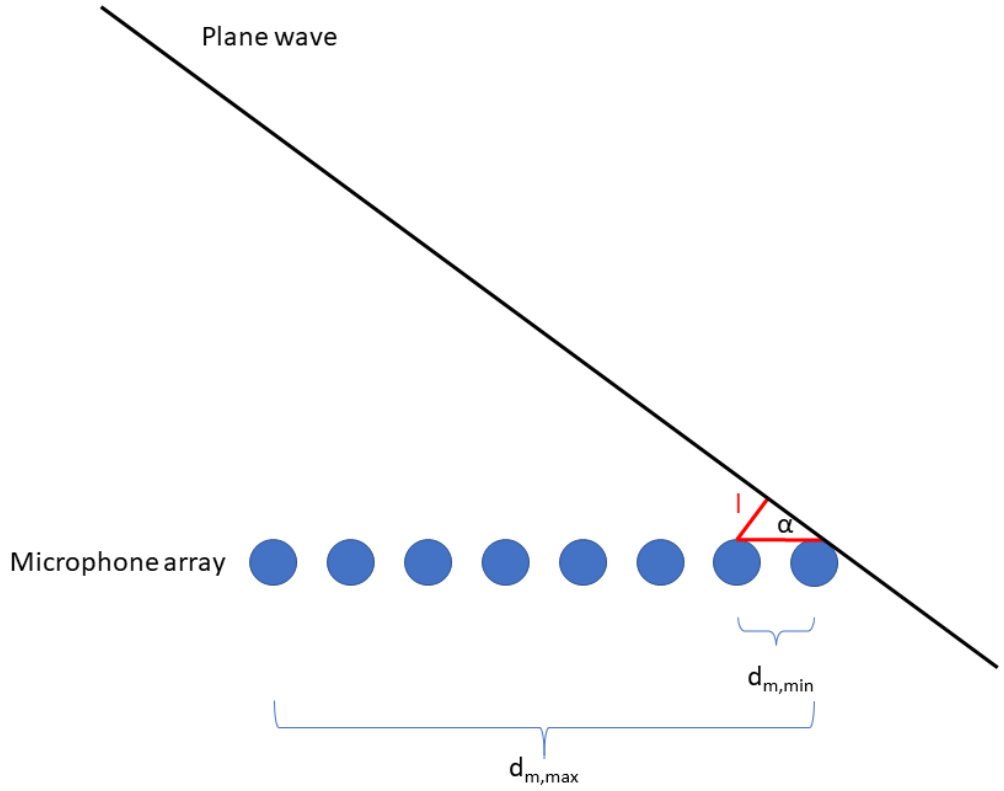


Figure 1.6: Schematic plot of a plane wave approaching a linear microphone array. The distances d_m between the microphones are known. The smallest distance $d_{m,min}$ defines the limit for the upper frequency of the array. The widest spacing $d_{m,max}$ determines the theoretical lower limit for the frequency of the array. The angle of incidence α of the plane wave on the array has to be estimated in order to define the direction of arrival of the signal.

vector. Contrariwise, *Capon*, belonging to the adaptive beamforming algorithms, readjusts explicitly to every condition. Further, *Clean SC* and *MUSIC* both appertain to the more complex deconvolution beamforming. This method consists of two parts: the first component creates an exploratory acoustic map with conventional beamforming, while the second component deletes phantom sources, leaving behind a very accurate image. The *EVOB* algorithm has been developed by the acoustic camera manufacturer *CAE* (Gütersloh, Germany) himself and thus very limited theoretical information is available for scientific investigation. In the next subsections 1.2.2, 1.2.3, 1.2.4, 1.2.5 a short insight into the theory of the algorithms is given. The correct functionality of each algorithm depends on the quality of the array calibration, but the requirements are not extraordinary.

1.2.2 Conventional Beamforming: Delay and Sum

Delay and Sum (DaS) beamforming is also called conventional beamforming and is one of the most applied beamforming algorithms [16]. Image 1.7 shows the theory behind the algorithm. The aim of DaS algorithm is to apply a specific delay Δ_m and a specific amplitude weight w_m to each microphone output, to sum the signals and achieve the highest constructive interference for an incident wavefront. When measuring with an array of sensors, in this case microphones, the incoming waveform read at each sensor accumulates [17]. Specifically, this means that signals arriving from different directions (different angles of incidence) encounter constructive or destructive addition (interference). An array with M microphones, each one having a specific position in space $\vec{x}_m = [x_m, y_m, z_m]$ is used to measure a wavefront $f(\vec{x}, t)$. As the microphone array does not represent a continuous plane, the waveform is spatially sampled with the sensors m, resulting in $y_m(t) = f(\vec{x}_m, t)$. In order to steer the antenna towards the source position the delays are adjusted in such a way that the captured signals by the microphone sensors add constructively [18]. The weights on the other hand are used to increase the shape of the main lobe and subsequently reduce side lobe levels (see figure 1.8). Since the *DaS* algorithm is not adaptive, sensor weights and delays are set and stored a priori for a predefined scanning plane and without knowledge of the incoming signal. In scanning mode, the array is steered in different directions in the scanning plane. When the direction of the microphone array coincides with the source, the maximum output power will be detected on the sensor array [16].

The time domain output of the DaS beamformer for a single point in the scanning plane can be written as:

$$z(t) = \sum_{m=0}^{M-1} w_m \cdot y_m(t - \Delta_m) \tag{1.7}$$

with M being the number of microphones in the array, w_m being the sensor weights and Δ_m being the delays. $y_m(t) = f(\vec{x}_m, t)$ is the measured signal at the m^{th} microphone sensor.
To be able to do so, an arbitrary set of listening points in the scanning plane need to be chosen a priori and these are defined as:

$$\vec{x}_s = [x_s, y_s, z_s] \tag{1.8}$$

The set of delays can then be calculated as (see figure 1.6):

$$\Delta_m = \frac{l}{c} = \frac{|\vec{x}_s - \vec{x}_m|}{c} = \frac{\sqrt{(x_s - x_m)^2 + (y_s - y_m)^2 + (z_s - z_m)^2}}{c} \tag{1.9}$$

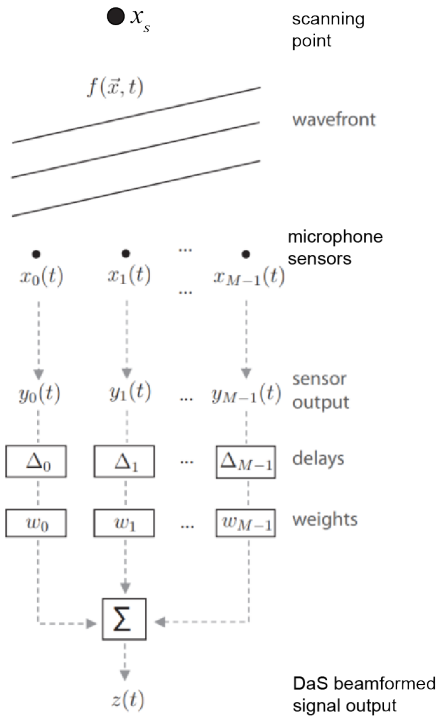


Figure 1.7: Theory of the Delay and Sum beamforming adapted from [16]. The wavefront is depicted as $f(\vec{x},t)$, the microphone sensors are marked as $x_0...x_{M-1}$. The electrical signal of each sensor is given as $y_0...y_{M-1}$. After adding the delays $\Delta_0...\Delta_{M-1}$ and the weights $w_0...w_{M-1}$ we receive the DaS output signal $z(t)$.

with c being the speed of sound. The reference point is in the centre of the acoustic camera. To define how responsive the array is to a single-frequency wave coming from any possible angle of approach we first consider the output of a single microphone:

$$\underline{y}_m(t) = e^{j(\omega t - \vec{k}\cdot\vec{x}_m)} \tag{1.10}$$

The wavenumber vector $\vec{k}$ gives information about the direction of arrival of the incident plane wave [15]. If the time delays now are established as phase shifts we obtain:

$$\underline{y}_m(t-\Delta_m) = \underline{y}_m(t)\cdot e^{-j\omega\Delta_m} \tag{1.11}$$

The signal $\underline{y}_m(t)$ will vary from sensor to sensor and $e^{-j\omega\Delta_m}$ is the associated phase shift for each microphone. The output equation 1.7 can be rewritten as:

$$\underline{z}(t) = \sum_{m=0}^{M-1} w_m \cdot \underline{y}_m(t)\cdot e^{-j\omega\Delta_m} \tag{1.12}$$

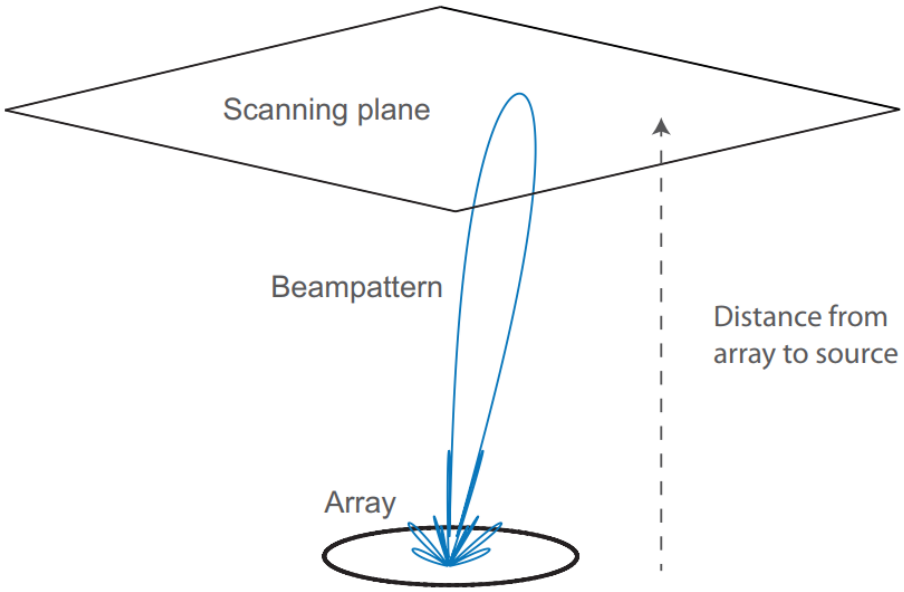


Figure 1.8: The concept of an acoustic camera. The output power of microphone sensors at multiple points on a scanning plane are calculated. Sensor signals of each scanning point are collected and then interpolated to create a valid image. The distance from the source to the sensor is predefined. The illustration is adapted from [16].

The delay for the correct "steering" of the vector is chosen as [1]:

$$\Delta_m = -\frac{\vec{k}}{w^\circ} \cdot \vec{x_m} \tag{1.13}$$

and inserting it in equation 1.11 results in a vector $\underline{\vec{y}}$:

$$\underline{\vec{y}} = \begin{bmatrix} \underline{y}_0(t) \cdot e^{j(\vec{k}\cdot\vec{x}_0)} \\ \underline{y}_1(t) \cdot e^{j(\vec{k}\cdot\vec{x}_1)} \\ \vdots \\ \underline{y}_{M-1}(t) \cdot e^{j(\vec{k}\cdot\vec{x}_{M-1})} \end{bmatrix} \tag{1.14}$$

[1] The reason for introduction and complete derivation of the parameter w° can be read in [16].

and $\vec{w}$ is the vector of the weights for each microphone:

$$\vec{w} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{M-1} \end{bmatrix} \tag{1.15}$$

The equation 1.12 can be rewritten in vector notation as:

$$\underline{z} = \vec{w}^T \underline{\vec{y}} \tag{1.16}$$

where $\underline{\vec{y}}$ represents the received signal of each microphone with its phase delay, $\vec{w}$ is the weight vector and T the matrix transposition [19].
Now, with the assumption, that the array has been steered to the desired direction it is possible to calculate the power of the output signal using vector notation:

$$\begin{aligned} P(\underline{z}) &= \mathbb{E}\{|\underline{z}|^2\} \\ &= \mathbb{E}\{(\vec{w}^T \underline{\vec{y}})(\vec{w}^T \underline{\vec{y}})^H\} \\ &= \mathbb{E}\{\vec{w}^T \underline{\vec{y}\vec{y}}^H \vec{w}\} \\ &= \vec{w}^T \mathbb{E}\{\underline{\vec{y}\vec{y}}^H\} \vec{w} \\ &= \vec{w}^T \mathbf{R} \vec{w} \end{aligned} \tag{1.17}$$

with $\mathbf{R} = \mathbb{E}\{\underline{\vec{y}\vec{y}}^H\}$ denoting the correlation function of the sensor outputs $\underline{\vec{y}}$ [16] and H the complex conjugate matrix transposition [19].

So the signal power of the *DaS* beamformer output for a single point is defined as:

$$\boxed{P(z) = \vec{w}^T \mathbf{R} \vec{w}} \tag{1.18}$$

1.2.3 Adaptive Beamforming: Capon

The Capon method was introduced as a novel signal processing technique in 2003 by Stoica et al. [20]. Compared to standard beamforming methods they show that their algorithm achieves higher spatial resolutions, is less sensitive to noise and interference, and can adapt to changes in the signal environment making it more robust in real-world applications. The adaptive part of the algorithm follows the *maximum likelihood estimation (MLE)*. After observing some data,

this statistical approach estimates the parameters of an assumed probability distribution. This is achieved by optimizing a likelihood function in a way that, according to the assumed statistical model, makes the occurrence of the observed data most plausible. The likelihood function depicts a general functional relation between a set of observed data and some unknown parameters.
For a known (defined) steering vector the Standard Capon beamforming (SCB) leads to very good results, it is data-dependent. For an unknown steering vector the results obtained with the Capon method can be worse than the ones resulting from a standard beamforming algorithm. To achieve an improved outcome, Stoika et al. [20] derived an improved algorithm known as the Robust Capon Beamforming (RCB). Their "robust" method is based on the theoretical foundation of the "standard" method, but takes into account the assumption that the knowledge of the direction of arrival can be inaccurate. The RCB can lead to more reliable results, but with the drawback of higher computational complexity. In the software of the used acoustic camera it was not specified which Capon algorithm was implemented.

1.2.4 Deconvolution Beamforming: Clean, Clean SC and MUSIC

CLEAN and CLEAN-SC

The *CLEAN* algorithm and its variant the *CLEAN-SC (Sparse Component)* algorithm are widely used beamforming techniques in astronomy and speech processing applications. The CLEAN algorithm, proposed by Högbom in 1974 [21], is an iterative algorithm that aims to remove the point spread function (PSF) from the observed data to recover the underlying sources. The point spread function is the response of a camera system to an impulse (a point source). Högbom assumes that the image consists of point sources, each of which can be represented by a Dirac delta function [22]. Often the synthetically obtained PSF's are not equivalent to real measured beam patterns and thus lead to inaccurate results. The CLEAN-SC algorithm [23] is an extension of the latter and allows to remove the side lobes of actually measured patterns. The algorithm starts by identifying the brightest point source in the image using conventional beamforming. It then iteratively subtracts the part of the source plot that exhibits spatial coherence with the highest peak source. This process is repeated, identifying the next brightest point source and subtracting it from the image until no point sources remain above a certain threshold.

The *CLEAN* algorithm is a widely used method. However, it has two critical limitations that should be considered. First, it always assumes point sources, which is not always the case. Second, it is computationally intensive, and convergence can be very time consuming.

MUSIC

The *MUSIC (Multiple Signal Classification)* algorithm utilizes eigenvalue decomposition to analyze signals [24]. The technique relies on the covariance matrix of the signals and can be used with both uniform and non-uniform linear arrays. It can be observed that, while the *MUSIC* algorithm is valid only if there is correlation among the signals from the same source across different microphones and no additional incoherent noise, the *MUSIC* algorithm can take into account the existence of incoherent noise and de facto employs it to identify the correct source position.

1.2.5 EVOB

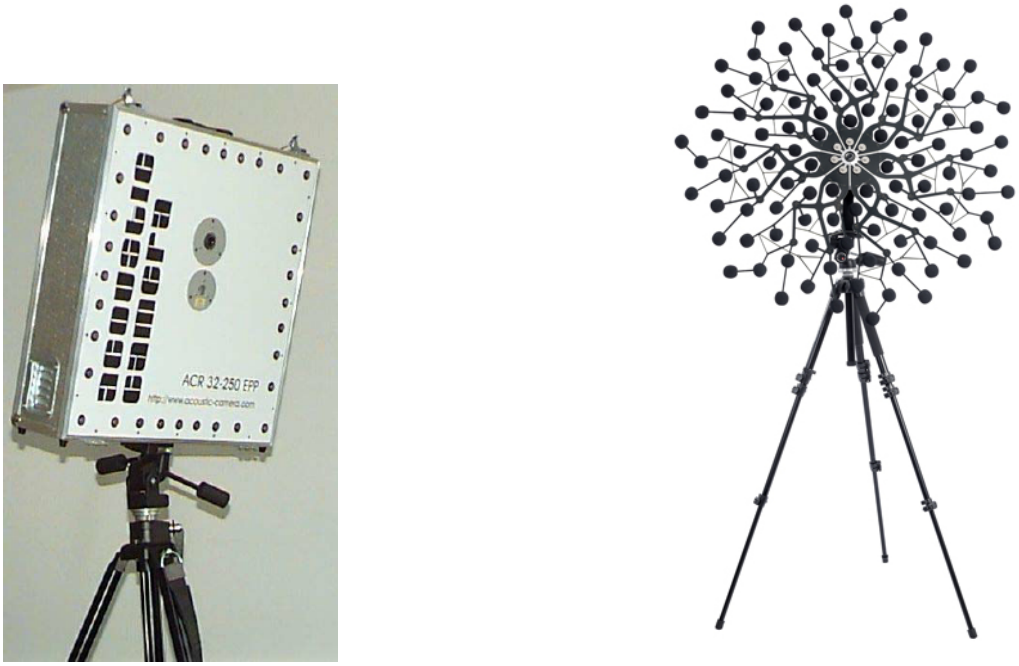
The *EVOB (EigenValue Optimized Beamforming)* is a refinement of functional beamforming implemented by the developer of the *CAE Systems* acoustic camera. The name derives from its utilization of matrix function mathematics [25]. Disappointingly, specific information regarding the exact implementation in the software is currently unavailable. However, theoretical proof is given that the use of functional beamforming can assure an increase in the flexibility of array design and broaden the dynamic range of the system.

1.3 Acoustic cameras

Nowadays acoustic cameras overlay optical photos taken with a digital camera with the corresponding acoustic colour map. This chapter will describe how this technique was developed, how it advanced to recent days and in which fields it is primarily used.

1.3.1 History of acoustic cameras

The first *listening device* which has been reported in a textbook by *Johnson & Dudgeon* was introduced during the first world war by the French army. The device was composed by a hexagonal array with six inverted horns for each ear. It was used to determine the direction of approaching air-crafts. Only in 1976 J. Billingsley published *"The acoustic telescope"* [14]. The first prototype line array worked offline and was tested with a Rolls-Royce Viper engine. The final system consisted of 14 condenser microphones and a mini-computer connected via 100 m long cables [26]. It was principally used to locate sound sources of jet engines. Developed to deliver rapid results it was compact enough to be transported in a car for field experiments. The analogue signal of the measurements could be saved on a magnetic tape recorder. Even though the maximum sampling frequency was 20 kHz, the data was stored on floppy disks (0.3 Mbytes of capacity), the images were displayed colour coded on a TV screen and the computer had a memory of 48 kilobytes [26], the structure was nearly identical to modern systems. Over the following years the rapid development of hard- and software resulted in increases of the sampling frequency and number of microphones, a higher digital resolution and reduction of computational time.
In 1999 GFaI (*Gesellschaft zur Förderung angewandter Informatik*) presented its first Acoustic Camera at the Hannover Messe [26]. It combined a microphone antenna with a digital camera. By overlaying the sound maps with a picture of the framed object almost in real time, the sound could be made "visible". This innovation surely impressed the audience and facilitated the further development and establishment of the microphone array technique for acoustic measurements.



(a) First acoustic camera, 1999.

(b) Modern acoustic camera, 2018.

Figure 1.9: Development of the array shape: substantial increase of number of microphones and resolution of camera.

1.3.2 State of the art of technology

Acoustic cameras developed along with the rapid increase of computer efficiency and data storage. Studies on the ideal array shape and size and data processing algorithms have been conducted for various applications. One of the first improvements was to increase the number of microphones. This action offers a more accurate source localization but the computational time (and cost) grows. Thus, different microphone distributions along the array can provide accuracy without increasing costs. The optimal geometry depends on the application fields of the array. Planar arrays are usually used for moving sources whereas spherical shaped arrays operate more in environmental matters. In 1998 Robert P. Dougherty registered a patent for a logarithmic spiral shaped microphone array. He declared that in this way the frequency range of the array would be wider because of the avoidance of spatial regularities [27]. However, to create 3D mapping, where information should be identical from all directions in space, spherical arrays are more adequate. In response to the increasing demand of acoustic cameras in different engineering fields many companies

have developed several products. *Gfaitech* proclaims three different sizes of ring arrays, a three-armed star-shaped array, a Fibonacci-arranged array and a self-design array kit for 2D sound imaging and two sizes of spherical arrays and two sizes of paddle arrays for 3D imaging. Every shape and size can be used with the same provided data recording system and analysis software [28]. *Sorama*, a dutch company, offers three distinct solutions. The portable *Cam iV64* with 64 microphones in a sunflower shape [29], the medium-sized *Cam64* with 64 microphones but a high resolution camera [30] and the *Cam1k* with 1024 microphones to assure far field beamforming down to 50 Hz [31]. *Seven Bel* [32] offers four different sound scanner for several applications. The functionality of their devices is distinct to the latter presented arrays. The scanner consist of eight microphones for the smallest diameter (12 cm) and five microphones for the three bigger diameters (50 cm, 132 cm, 254 cm) mounted on a rotating arm which allows to simulate a much higher number of sensors. According to the manufacturer, each product has three different rotation speeds, making it suitable for different frequency ranges. *Seven Bel* declares their biggest antenna as the largest on the world market and a performance range from 125 Hz to 4000 Hz. An additional advantage is the use through a simple mobile application (Android only). The data is transmitted to the companies' servers, processed and the results are sent back to the user within a very short time.
A very dissimilar device is the *Norsonic Hextile* acoustic camera [33]. The producer developed one single *Hextile* with 128 MEMS microphones that can be combined with further Hextiles to improve the resolution at lower frequencies. Each tile weighs only 3 kg and has a diameter of 46 cm. Combining three tiles together results in a 96 cm diameter and allows a lower frequency limit of 220 Hz (410 Hz for the single tile). Moreover, the three tiles can be attached to a larger frame with a diameter of 146 cm, creating the low frequency mode configuration with a lower limit 120 Hz. The three dispositions can be seen in picture 1.10. However, one of the biggest differences from the other manufacturers is that the post-processing analysis software runs on Mac systems, not Windows.

Another well-known producer of acoustical measurement equipment, who offers an acoustic camera is *Brüel & Kjaer* [34]. Their solution is a handheld wheel array (diameter 35 cm) with 30 randomly placed microphones. They use a self produced quarter inch microphone with a frequency range from 50 Hz to 20 kHz. However, the company does not specify the acoustic camera's frequency range in the product data sheet. To visualize and post-process the acquired data four distinct software systems have to be installed on a windows system. There is also the possibility to use the array with an android tablet. Brüel &

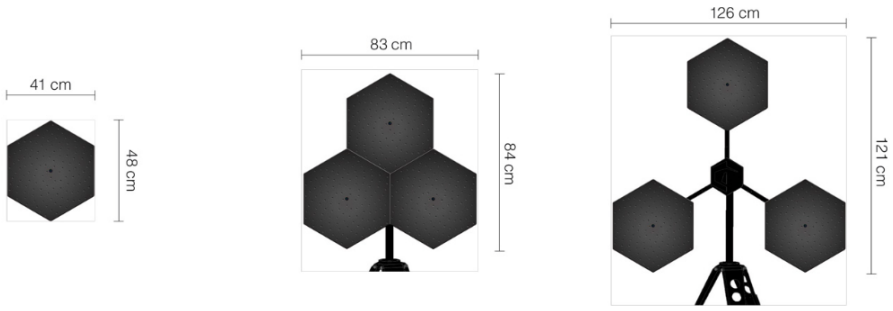
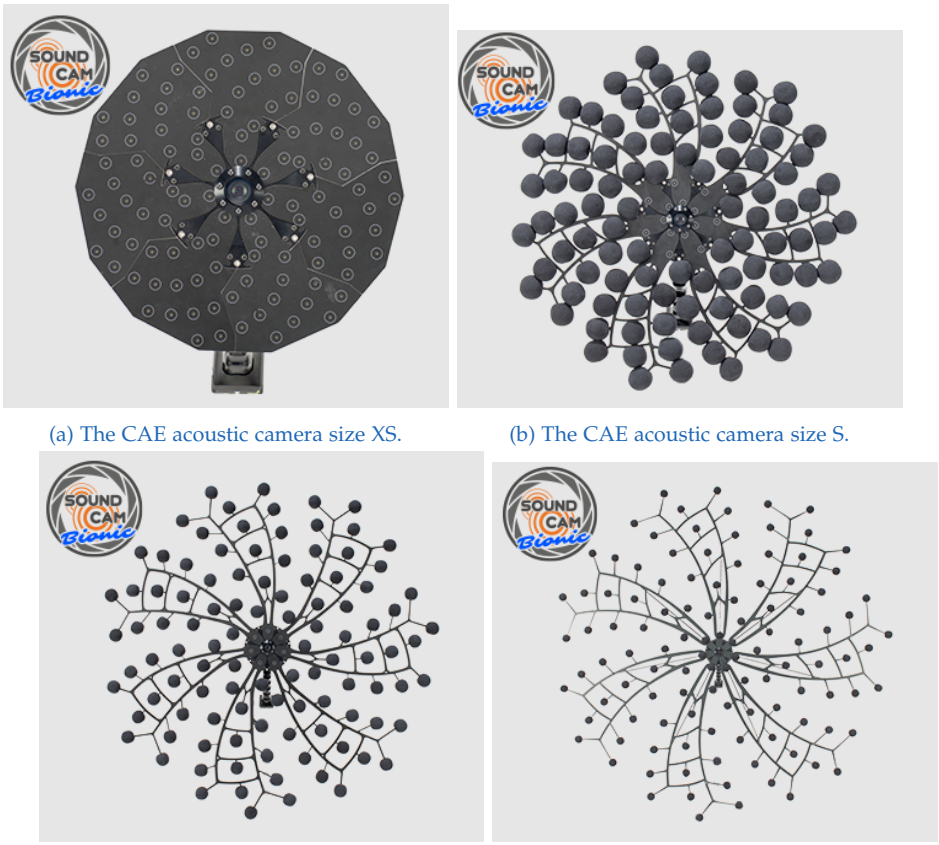


Figure 1.10: Different comibinations of the *Norsonic Hextile*.
left: single tile, *centre*: multitile, *right*: multitile low frequency mode.

Kjaer presents their product as very effective especially in small spaces like cars, airplane cockpits or machinery rooms. In the specific data sheet the two ways to create an acoustic map are mentioned: beamforming or holography. No specification about different beamforming-algorithms is given.
The last producer presented in this work is *CAE - Software and Systems* with five different *SoundCams* [35]. They offer three different hand-held solutions and four larger microphone arrays. The arrays measure 28 cm, 54 cm, 100 cm and 170 cm respectively. Every acoustic camera runs with the *CAE Noise Inspector Software* and an analysis with six different beamforming algorithms is possible. The software can be used in online mode, for fast real-time results or in offline mode, for more accurate frame-by-frame results. The different array sizes allow measurements in different frequency bands and subsequently for different applications. Each antenna persists of 112 MEMS microphones arranged in a Fibonacci-spiral around an optical camera, merely the allocation changes. Picture 1.11 shows the different acoustic camera arrays. For the measurements in the port of Livorno the University of Pisa (Italy) kindly provided the CAE SoundCam Bionic L.

The latter mentioned producer declare their product as the best state of the art technology for various reasons. The most important benefit is the simple operation without prior training and with few steps to the final result [32], [35], [34]. For the hardware an often mentioned specification is the light weight and compactness of the arrays and consequently the hand-held operation possibility [35], [34], [33], [32]. *CAE* [35] specifically states their acoustic camera size L with the diameter of 1.7 m and weight of 5 kg as a hand-held option.
The third well promoted feature is the fast data processing partially with simple mobile phone applications [32], [35]. Again, *CAE* [35] assure a variety of algorithm for high resolution images at the most favorable price. The main problem that arises are the closed source algorithms implemented in the software. These closed-source design do not enhance the usability of the acoustic



(a) The CAE acoustic camera size XS.

(b) The CAE acoustic camera size S.

(c) The CAE acoustic camera size M.

(d) The CAE acoustic camera size L.

Figure 1.11: Four different sizes of the CAE SoundCam Bionic. From a) the smallest size with a diameter of 28 cm, b) with a diameter of 54 cm, c) with a diameter of 100 cm to d) the largest size with a diameter of 170 cm. The number of microphones in the array is 112 for each size.

cameras for research purposes. Further, they do not give the possibility to fully understand their range of capabilities and subsequently do not provide the scientific credibility.

1.3.3 Application fields

The different sizes and consequently the different working frequency range makes the acoustic camera a versatile tool for sound recognition. It can be used in machinery construction to recognize noisy elements and dominant sound sources within a production plant [32]. Another possible appliance is in the building acoustics field. Sound leaks from walls, windows, doors or ventilation

systems can be detected and visualized. Further, internal and external noises of vehicles can be tracked down. The acoustic behaviour of tires and wheels or the engine can be observed and compared to simulation models. Disturbing vibration within the driver's cabin or leaks which allow noise to penetrate into the vehicle can be localized with the help of the smaller acoustic cameras. Bigger and thus cameras with a wider working range are often used for environmental issues. Long mechanical production lines with multiple conveyors, cranes, pipes, pumps emit dissimilar sounds which can be recognized and categorized with an acoustic camera. The larger acoustic cameras with a diameter of 5m (*wind turbine* of CAE Systems [35]) allow to create images of sources with frequencies down to 50 Hz. This application especially comes into use when measuring and displaying noise from wind turbines or big transformers within electricity networks.
Since there are different manufacturers of acoustic cameras and therefore also products with different specifications, the areas of application are diverse and variable.

1.4 Aims and Objectives

The objective of this study is to enhance the comprehension of the microphone array and its associated software analysis tool, by conducting a performance benchmarking experiment in a controlled environment, namely a free field. The subsequent step involved the development of an image analysis algorithm to assess the precision of each beamforming algorithm utilized, with regards to the angle of incidence of the sound, distance between the sound source and the acoustic camera, selected algorithm, and the evaluated third-octave frequency band. The accuracy of the algorithms was analyzed in the context of these variables, in order to achieve a better understanding of the array's overall functionality. This approach was designed to provide insights into the strengths and limitations of the system, as well as to identify opportunities for improvement. By utilizing this methodology, the study aimed to generate open-source data that can contribute to optimize the array and software analysis tool, ultimately enhancing the accuracy and efficiency of acoustic imaging in real-world applications. This approach can significantly enhance the precision of the noise maps generated within the INTERREG Maritime Programme. This optimization-based methodology offers an objective and data-driven approach to noise mapping, which can lead to improved decision-making and mitigation strategies for noise pollution, while simultaneously create comprehensible maps for non-acoustic specialists.

2 Materials and Methods

The team spent various measurement days at the port of Livorno on behalf of the *INTERREG* project. The aim was to measure different situations of continuous ambient noise with a sound level meter and the acoustic camera. The obtained noise levels should then be used to update the noise maps of the area. However, when the first data was to be analysed with the *CAE acoustic camera* there was no certainty about the influence of various parameters on the final result. This parameters being the distance from the source, the incidence angle of the sound wave and the used algorithm for analysis. In order to be able to validate the effect of these parameters on the deviation between the real source location and the calculated one, the idea of an experiment with ideal conditions was born. As mentioned in chapter 1.1 the port is an acoustically complex area. Many different noises emerge simultaneously from diverse, often unknown sources. For this reason the eagerness to conduct measurements in a silent area with controlled noise source conditions arose.

A whole process from the idea, over the measurement and data handling to the final results had to be developed. The result of the final workflow can be seen in figure 2.1 on page 23. The key questions were: how can be measured and under which conditions. As the main aim was to study the functionality of the algorithms in the *Noise Inspector* software, a way to objectively compare the algorithms had to be created. The measurement in an isolated free field with low background noise seemed a legitimate choice. To ensure an admissible signal to noise ratio for all four distances the measurement was conducted in parallel with the acoustic camera and a sound level meter at the same position. In the first post analysis the acquired data was evaluated with the software *dBTrait* and the applicable results selected for closer investigation. The next step was to extract and save the images from *Noise Inspector*. In this step the same procedure was applied to each measurement data:

- The same time-frame was selected (1 sec);
- The same third frequency bands steps from 250Hz to 4000Hz were applied;
- The same image resolution was chosen.

The created images were then exported to *Photoshop* and equally cropped to the size of 1000 x 575 pixel. At this point the image processing and algorithm

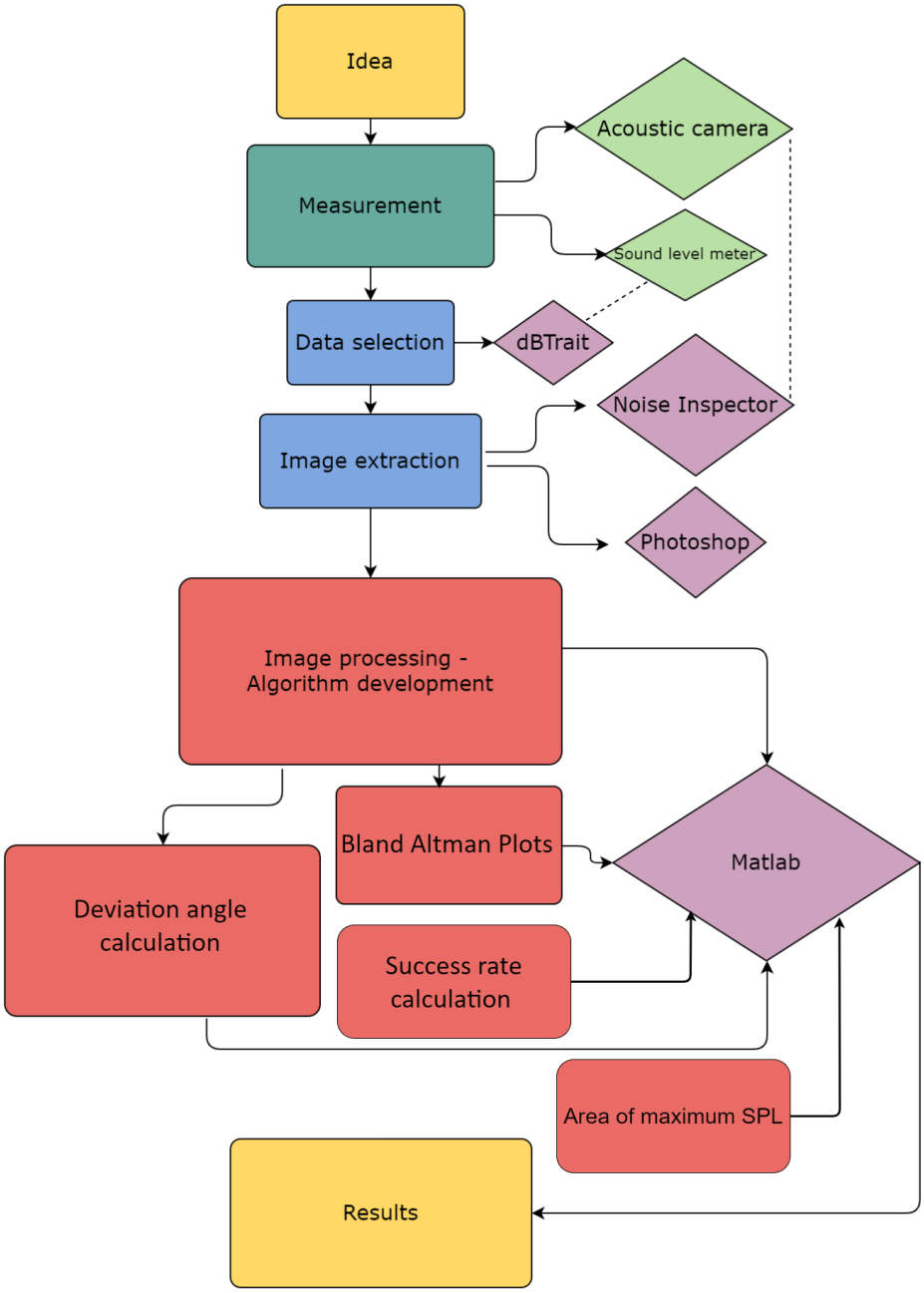


Figure 2.1: Flowchart of the whole process.

development began. Four different analysis were conducted to assure the comparability of the measuring method with the acoustic camera vs. the measuring

method with the naked eye. First, the success rate within the measurement for each algorithm was calculated. In a further step, the geometric analysis with Bland Altman plots and the deviation angle was calculated. The last analysis consisted in calculating the area of maximum sound pressure level in the images. The whole image processing and algorithm development was build up in *Matlab*. The final graphic results were then analysed and compared. Each step of the workflow is described in the following sections.

2.1 Experimental setup

The experiment was conducted in an abandoned field in the Tuscan countryside far away from car traffic or population. The only disturbing sources where some airplanes overflying the area and some singing cicadas. However, it was paid attention to ensure a high enough signal to noise ratio during the measurements at every distance. With a sound pressure level meter we ensured a constant signal-to-noise level of at least 12 dB during each measurement. Only a few light bushes were in the area and the next house was about 80 m from the acoustic camera position. The smallest distance between the source and the house was 20 m. The ground was fully covered by grass and thus assured an absorbing ground. Hence the field provided (an almost) reflection free environment. Image 2.2 shows the acoustic camera during the measurements in the field.
The acoustic camera had a constant position (height 1.2 m), whereas the noise emitting source was moved to 12 different positions in the field at a height of 1.2 m. The used source was a *Norsonic Dodecahedron Nor229* emitting white noise generated with a noise generator and amplified with the *Norsonic power amplifier 260*. At every point a signal of the duration of 10*s* was emitted and simultaneously acquired with the acoustic camera and a sound level meter. The exact setup can be seen in table 2.1 and figure 2.3.

Figure 2.2: Acoustic camera in the free field setup before conducting the performance test under the different conditions.

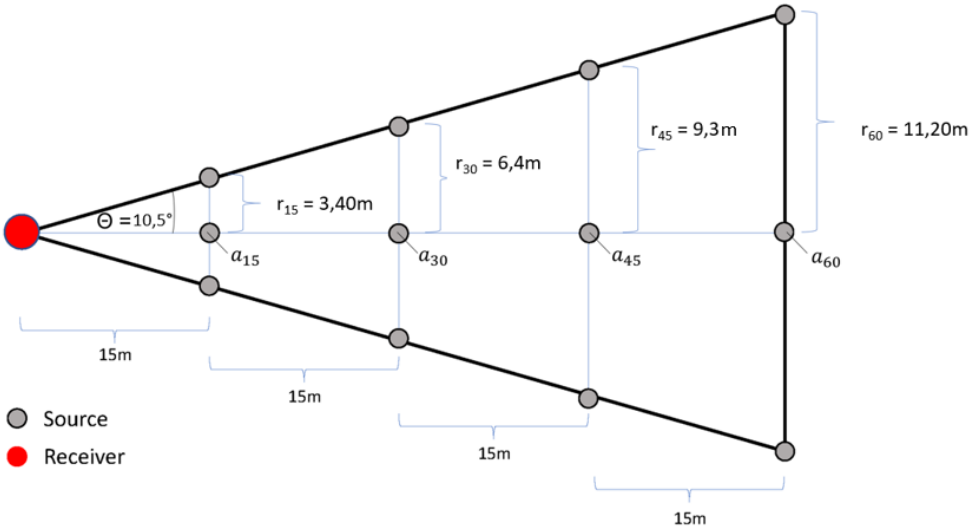


Figure 2.3: Measurement setup in the field.

Table 2.1: Positions of the source in respect to the acoustic camera.

distance a [m]	angle θ [°]	orientation
15	0	-
15	10,5	left
15	10,5	right
30	0	-
30	10,5	left
30	10,5	right
45	0	-
45	10,5	left
45	10,5	right
60	0	-
60	10,5	left
60	10,5	right

2.2 Data selection and image extraction

After completion of the measures the right data had to be extracted for the successive image analysis. A first analysis was done by checking the Signal to Noise ratio for each distance. This analysis was made by looking at the sound level meter signal with the software *dBTrait*. This software was chosen because it is the noise data processing software for the *Fusion* sound level meter used for the measurement.
Another key factor was the minimum frequency functionality of the acoustic camera array. With the formula 1.6 described in chapter 1.2.1 on page 7 the theoretical lower frequency limit for the array with a diameter of 1.7m was calculated.

$$f_{min} = \frac{c}{2d_{m,max}} = \frac{340[\frac{m}{s}]}{2 \cdot 1,7[m]} = 100[Hz] \tag{2.1}$$

However, for a higher resolution it is recommended to double the theoretical value [36]. This value was decisive in the selection of the frequency range to be considered.
In the next step the acoustic camera software *Noise Inspector* came into use. This software allows to select specific time frames from the measurement files. From this frames the user can either produce a video, with flexible frame rates or a static image that shows the result for the time averaged data. Further it is possible to choose if the SPL should remain *linear* or *A-weighted*. There is also the

possibility to introduce the distance at which the acoustic camera was placed from a (possible) source. Considering that the software always works from a receiver perspective, the user sometimes does not know the exact position of the source he is measuring. For this case though, the distance was known. The software offers the option to choose between 8 algorithms. For this data analysis only the five most relevant were used:

- Delay and Sum (DAS)
- Capon
- Clean SC
- EVOB
- MUSIC

Each algorithm can be applied to a desired frame at a specific frequency value. It is possible to select a single frequency, a third-octave frequency band or an octave frequency band. In this case results for the third-octave frequencies from 250 Hz up to 4000 Hz were produced. Hence, for every position of the source an image averaged over 1s for each third band frequency within the range from 250 Hz to 4000 Hz and for 5 different algorithms was created. The result were 780 images to evaluate. The images can easily be exported and saved as *.png* files.

$$12 \text{ source pos.} \cdot 13 \text{ third octave freq.} \cdot 5 \text{ algorithms} = 780 \text{ images} \tag{2.2}$$

Before passing on to the image analysis the images had to be cropped to the same size and the true source location determined. The exact pixel location in x and y direction for each source position was established in *Adobe Photoshop* and then transcribed into an excel sheet. These coordinates will later be compared to the results of the algorithm calculation and give an outlook about the accuracy of the antenna varying with the specified parameters.

2.3 Image processing - Algorithm development

The main purpose of the image processing was to detect the **calculated** source centre position and compare it to the **real** source centre position. The created images show a heat map around the highest, from the acoustic camera perceived sound (SPL-level). Throughout the measurement in the field we assured a sufficient SNR by constantly controlling it on the sound level meter, but we did not write down the exact level. Previous measurements in the port simultaneously with the acoustic camera and the sound level meter showed, that the value measured with the acoustic camera and the value measured with the sound level meter are not equivalent. Hence, for the algorithms that give the information (*Capon* for example cannot give the information as it is visible in figure 2.4), the results of the absolute *maxSPL* value are not reliable.
In picture 2.4 each colour means a $1dB$ reduction. Thus, the coordinates of the *red Part* of the heat map had to be identified. The implementation of the whole image segmentation and analysis was developed in *Matlab*.

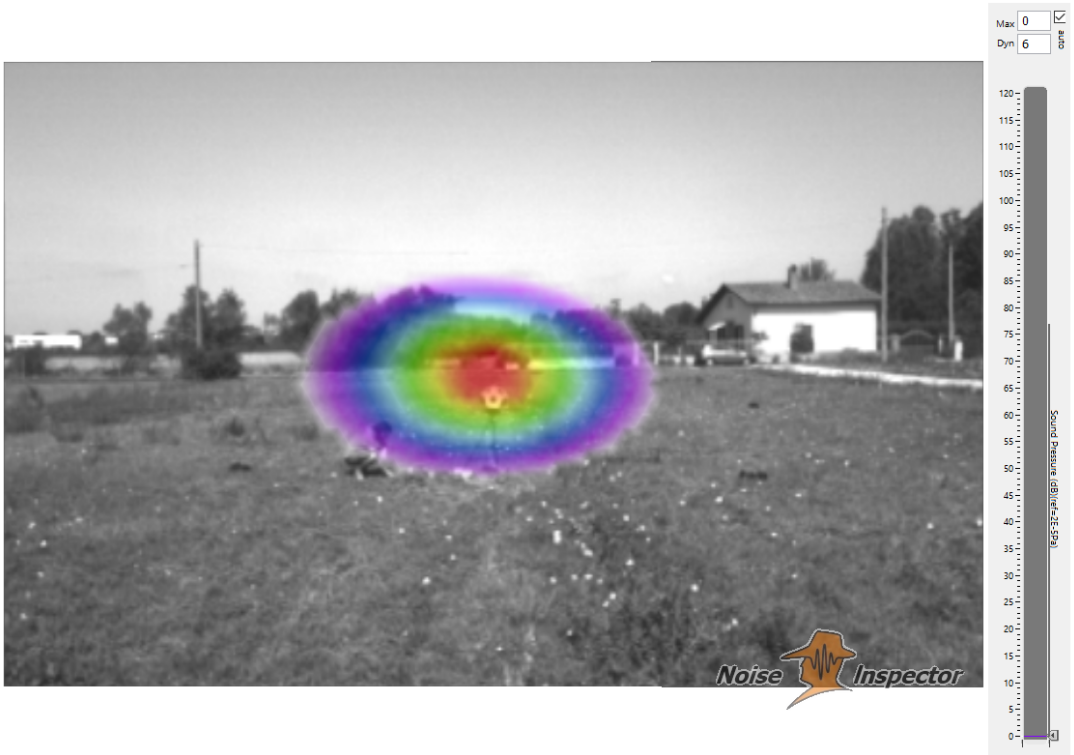


Figure 2.4: Original image evaluated with Capon algorithm at 315Hz.

filename	**functionality**
loader.m (function)	loads right files in directory
natsortfiles.m (function)	sorts filenames alphabetically and from low to high numbers
imageprocessing.m	standardise image size, 2 step filter to isolate red part of image, identify absolute centre of source, exclusion of faulty images, creation of Bland Altman Diagrams, create comparison images with algorithm vs. groundtruth (see image 2.6)
horizontal_angle_deviation.m	calculation and representation of the horizontal angle deviation from the real source position
vertical_angle_deviation.m	calculation and representation of the vertical angle deviation from the real source position
halo.m	creates graphs of area of maximum SPL

The first file to use is *imageprocessing.m*. In this file the first variable to define is the path for the access to the images created with *Noise Inspector*. It is important, that all the images have the same name composition. For this work they started with the position of the source, then the used algorithm and finally the third frequency band separated by an underline, resulting in for example *c15_capon_250*. Once the search path is defined the *loader* function is called and within the loader function the *natsortfiles* function to sort the filenames alphabetically and from low to high numbers. With this sorting function it was possible to ensure that the images were always processed in the same order and thus their position in the created structure array was always predictable. In the *imageprocessing* file a large data structure array, where the name of every image, its size (784 x 1152), its cropped size (576 x 1001), the image information itself, the pixel coordinates of the coloured map and every further calculation is stored, is created. Also the Bland Altman calculation and graph plotting is done in this file.
In *horizontal_angle_deviation* and *vertical_angle_deviation* the angular deviation from the real source position is calculated. A more accurate explanation of the computational variables can be read in chapter 2.6.
Finally, in the *halo* file the detected area of maximum SPL level is calculated and plotted. One graph for each source position was created, allowing a comparison of the single algorithms. Also, the real size of the source was included in the graph as a constant line to facilitate the interpretation of the results.
In the next paragraphs a more detailed explanation of all the calculations is given.

Firstly, the images were divided by distance from the source (15m, 30m, 45m and 60m), were loaded into the software and cropped all to the same size. Then the first step of the image processing was built up. The pictures were transformed from the *RGB format* to *HSV format*. *RGB* stands for *red, green, blue*

while *HSV* stands for *hue, saturation, value*. The first value corresponds to a value on a colour wheel starting from red to orange, yellow, green, cyan, blue, magenta and back to red. Saturation goes from 0 to 1, where 1 is the maximum saturation. The third value indicates which color amongst red, green and blue has the maximum value. The transformation into this format allowed to isolate the maximum of the recognized source and thus the centre of the created heat map. This first solution for image segmentation is visible in figure 2.5. First attempts to filter the red colour from all the images, led to imprecise results. Without transforming the RGB-image to a HSV-image the right red value could not be isolated correctly for all images. In many cases the purple colour of the heat map would be considered as similar enough to the red one and subsequently distort the result. A new solution had to be thought of. On this wise the idea with the HSV segmentation came up. The results were very satisfying.

Figure 2.5: Image after first step of filter: HSV convertion and filtering.

The second step consisted in isolating the red centre and transform it into a binary mask. The images were transformed from HSV-values to LAB-values. LAB is an alternative way of colour coding and the combination of both, the HSV- and the LAB-filter lead to the most accurate results. The *L* stands for *Light*, the *A* gives information about the *green* and *red* value and the *B* carries the information about the *blue* and *yellow* value in the single pixel. Then the remaining red value was transformed to a grey one. This change allowed to

reduce the 3D matrix to a 1D matrix allowing to calculate the centre of the source for each picture. By creating histograms of the set bins in the images, the maximum width, given with the number of set pixels in the x-axis and the maximum height, given with the number of set pixels in the y-axis were known. Assuming an elliptical shape for the red area, the intersection of these two lengths gave the centre point of the red area.
Two filters for `errorcases` were introduced. The first one excludes all centre points placed at the borders of the image. That means the point contains either the maximum or the minimum pixel in the x-value or the y-value. The second case excludes all images were the algorithms did not detect any source and thus would place the centre at the top left corner.
An example of an excluded result can be seen in figure 2.6. The image shows

Table 2.2: This table shows how many images for each position of the source were marked as valid after excluding non valid images.

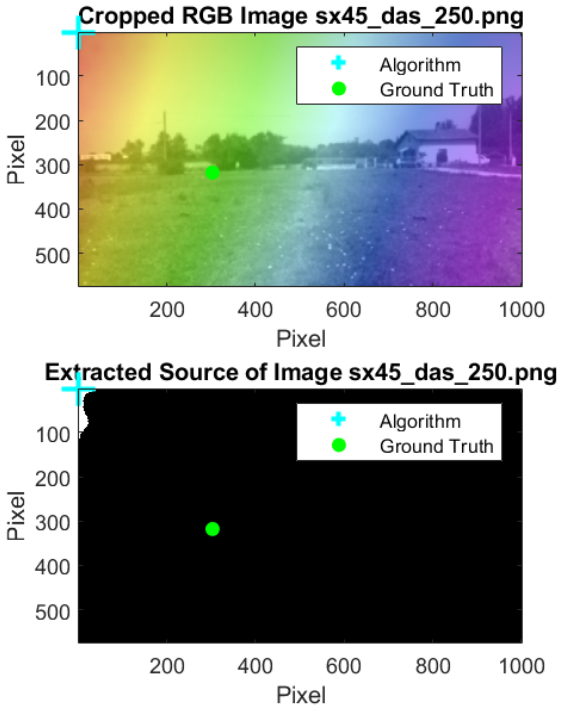
measurement	No. original files	No. valid files
15m_centre	65	64
15m_left		57
15m_right		60
30m_centre		63
30m_left		57
30m_right		59
45m_centre		63
45m_left		56
45m_right		56
60m_centre		62
60m_left		58
60m_right		59

the evaluation of the left position of the source at a distance of 45 m (*sx45*) with the *DaS* algorithm for the *third frequency band* of 250 Hz.
Whereas figure 2.7 shows a good placement of the source.
The image shows the evaluation of the centre position of the source at a distance of 15m (*c15*) with the *EVOB* algorithm for the *third frequency band* of 250 Hz.
The table 2.3 shows the abbreviations used in the title of the images.

Table 2.3: Explanation of the syntax in the name of the images.

Abbreviation	Meaning
c	centre
sx	left
ds	right

(a)



(b)

Figure 2.6: Example of wrong results evaluated at the left position of the source (sx) at a distance of 45 m. Image (a) shows a non valid image which was evaluated with the Clean SC algorithm at the third octave frequency band of 400 Hz.
Image (b) shows a different non valid image created with the Delay and Sum algorithm (das) in the third frequency band of 250 Hz. The source's centre (turquoise cross) is placed at the top left corner.

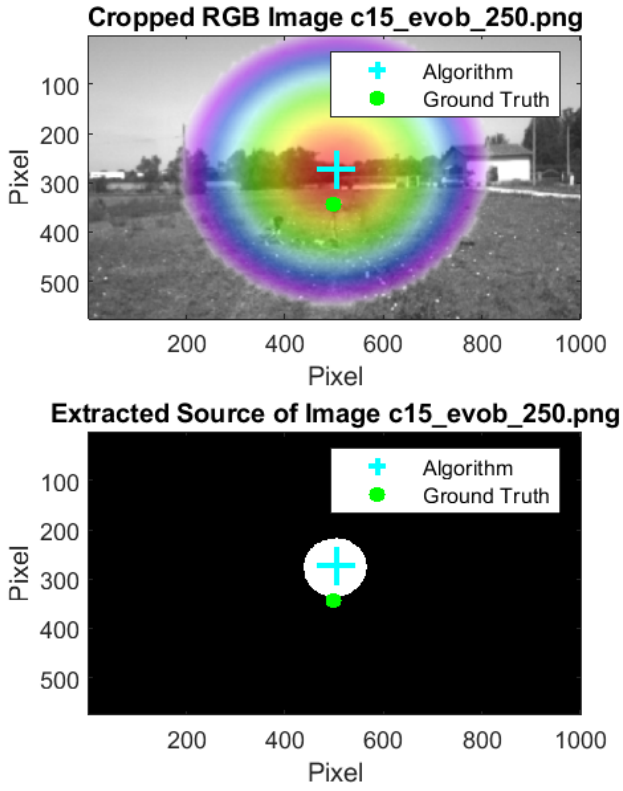


Figure 2.7: Example of a good result evaluated at the centre position of the source (c) at a distance of 15 m with the Evob algorithm in the third frequency band of 250 Hz. The algorithm places the centre of the source close to the declared real source position.

2.4 Success rate calculation

The first calculated value was the success rate for each algorithm. It defines which one is the most reliable in terms of source identification outcome. Every algorithm should produce a solid result in every third octave frequency band from 250Hz to 4000Hz. Hence generates 13 valid values for each source alignment (*left*, *centre*, *right*) at every distance (15 m, 30 m, 45 m, 60 m). To calculate the success rate the three source alignments were merged together for each algorithm, but kept separate for each distance. So one algorithm should recognise 39 valid source points at each distance. Consequently, the value 39 is equivalent to 100% success rate.
The images were filtered in two ways: firstly, if the algorithm places the source at the border of the image and thus the distance between the real source position and the calculated source position gets too big, the image gets excluded from the assay. Secondly, if the algorithm does not recognise any source in the image. However, to avoid having different sized vectors and thus making the data handling very complicated, the distance value of the image was not deleted but set to NaN. The success rate calculation first counts the NaN values, then subtracts them from the 39 possible values and finally calculates the percentage of success. As an example: the *Capon* algorithm produced 34 valid images resulting in a percentage of success of 87%. So the *percentage of success*-graph shows 4 groups (the four distances) of 5 bars (the five algorithms).

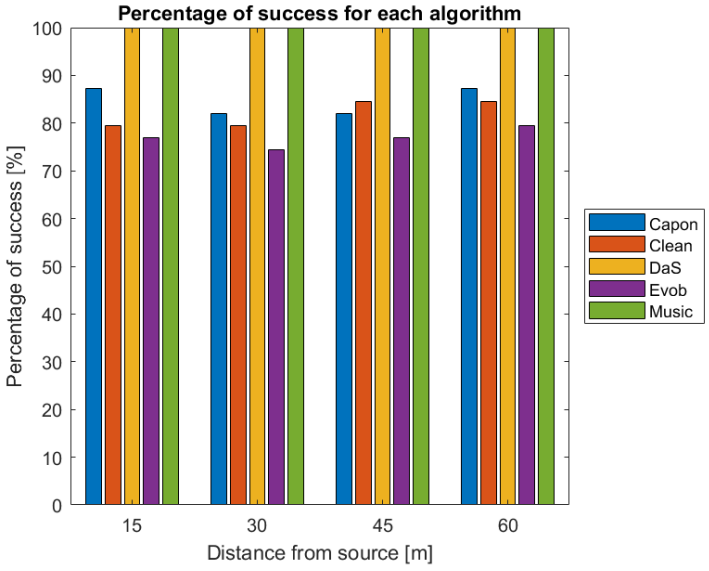


Figure 2.8: Percentage of success for all the distances.

2.5 Bland Altman diagrams

Once all images were filtered, a *Bland Altman Plot* for the calculated centre and the ground-truth centre were created. This method is mainly used in medical environment to compare a new measurement method with a golden standard. J. Martin Bland and Douglas G. Altman developed it to analyse the agreement between two different assays [37]. For two data sets S_1 and S_2 the plot is created as follows:

$$S(x,y) = (\frac{S_1 + S_2}{2}, S_1 - S_2) \tag{2.3}$$

In addition, on the y-axis the mean value of the difference between the assays and the positive and negative limits of agreement (LOA) are plotted. The LOAs are calculated as $d - 1.96s$ and $d + 1.96s$. With d being the mean difference between the samples and s the standard deviation of the differences. The limits of agreement of 95 % for each comparison were chosen according to common practice. The mean value of the difference gives an insight into how large the deviation between the assays is.
The Bland Altman analysis was elected over the correlation coefficient measure because it relieves the true agreement between assays. Whereas the correlation coefficient r only gives information about the strength of the relation between two data sets. Two sets which are strongly correlated can still have poor agreement. This was the key point in favour of creating plots for all distances and directions with the Bland Altman method. To do so, the assumption that the differences between the measurement variables are normally distributed was made.
In this work S_1 are the source points defined by the observer, also called 'ground-truth' and S_2 are the source points detected by the different algorithms. For better readability of the graphs the results for different algorithms were marked with different colours. Thus there are 13 points in each graph with equivalent colour representing the 13 third octave frequency bands. The algorithms with the corresponding colours are listed in table 2.4. A Bland Altman graph for each

Table 2.4: Legend of colours for each algorithm in the Bland Altman plots.

Algorithm	**colour**
Capon	blue
Clean SC	orange
DaS	yellow
Evob	purple
Music	green

distance, direction and x- and y-axis was created. The result were four different groups of graphs consisting of six graphs each.
Figure 2.9 shows an example of a Bland Altman diagram. The points lay primarily on a straight line because the value for S_1 remained the same for the one specific source position while the value S_2 fluctuated. A changing value also in the assay S_1 would lead to a more scattered distribution of the points. The line has a negative gradient because with an increasing value of S_2 the difference between the assays $S_1 - S_2$ decreases, while the mean value $\frac{S_1+S_2}{2}$ increases.
Figure 2.10 shows the images which lead to the two marked points in the diagram. On the x-axis the algorithm places the centre of the source twice as far as on the y-axis (~350px vs. ~100px). Thus, the result for this source setup in combination with the *Capon* algorithm leads to a considerable difference between the assay S_1 and S_2.

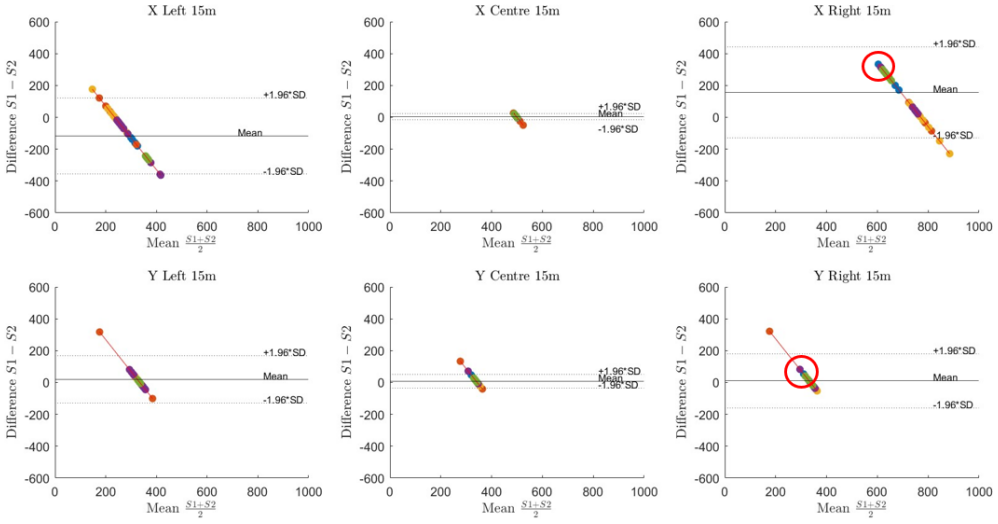


Figure 2.9: Bland Altman diagrams for 15 m distance to the source. The graphs show the difference between the *ground truth* of the source position S_1 and the *calculated* source position S_2 in pixel over the mean value between the two assays S_1 and S_2 in pixel. The two red circles mark the points which result from the images in figure 2.10.

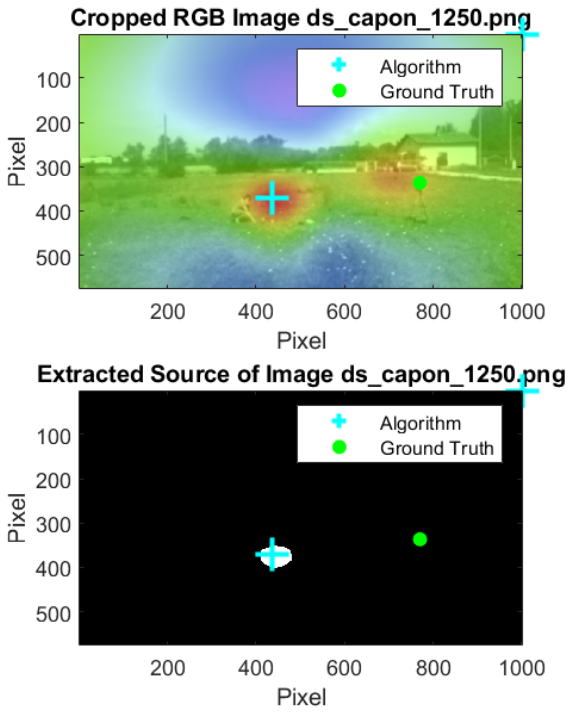


Figure 2.10: These two images show the result at the third-frequency band of 1250 Hz for the analysis of an image with the *Capon* algorithm with the source at the distance of 15 m and in right position.

2.6 Deviation angle calculation

Horizontal deviation angle φ_h

Image 2.11 shows the underlying concept for the calculation of the horizontal deviation angle φ_h between the real source side position $S_L(left)$ and the one detected by the different algorithms S_A. R stands for the receiver, in this case the centre of the acoustic camera; a is the distance from the receiver R to the central position of the source S_c; r_x is the distance between the central position of the source S_c and the real side position of the source $S_L(left)$; b_x is the distance between the real central position of the source S_c and the calculated side position of the source S_A. All these values are still in pixel. The transformation to meter only happens in a second moment with the calculation of the angle deviation error. The angles Θ_h and α_h are calculated in radian, but transformed to degrees in a second step. So the final result φ_h and its error $\Delta\varphi_h$ is a horizontal deviation angle from the real source position.

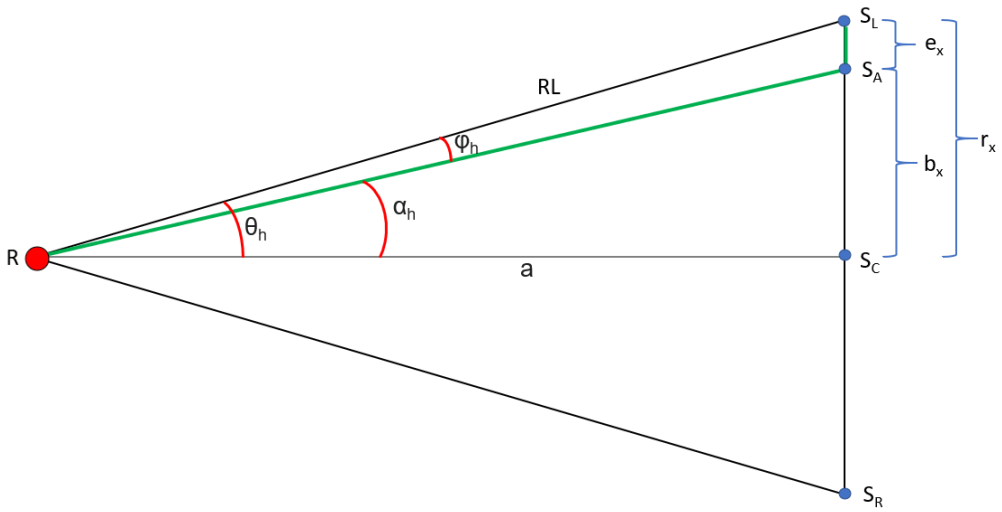


Figure 2.11: Scheme for the horizontal deviation angle φ_h calculation.

Since the real size of the dodecahedron was known it was possible to transform all distances in the picture from a **pixel**-value to a **meter**-value and subsequently calculate the **deviation angle** φ_h and its **error** $\Delta\varphi_h$ for each algorithm for every setup.

All the error equations were looked up in [38].

Before getting to the final result, numerous steps were necessary. Firstly, the transfer function f_T from pixel to metric for each distance had to be calculated. The [known] size of the source S_S divided by the number of pixel in the picture N_{PS} taken with the acoustic camera results in:

$$f_T = \frac{S_s}{N_{PS}} \tag{2.4}$$

$$\Delta f_T = \sqrt{\frac{\Delta S_s^2}{N_{PS}^2} + \frac{S_s^2}{N_{PS}^4} \cdot \Delta N_{PS}^2} \tag{2.5}$$

The error for the transfer function Δf_T was very small since ΔS_s was only $\pm 0.01m$ [information from the datasheet of the source] and ΔN_{PS} was assumed to be only $\pm 1 pixel$.

Table 2.5: Multiplying coefficient for each distance from the acoustic camera to transform pixel into m.

distance a **[m]**	**source size** S_S **[m]**	**source size** N_{PS} **[pixel]**	f_T $[\frac{m}{Pixel}]$
15	0,42	32	0.0129
30	0,42	17	0.0242
45	0,42	12	0.0350
60	0,42	10	0.0420

Table 2.5 shows the transfer function coefficient for each distance.
With the transfer function coefficient f_T and the distance between the real side source point S_L and the calculated side source point S_A in pixel N_p it was possible to calculate the **deviation** e_x from the real source position to the calculated source position in $[meters]$ and its error Δe_x.

$$e_x = N_p \cdot f_T \tag{2.6}$$

$$\Delta e_x = \sqrt{\Delta N_p^2 \cdot f_T^2 + \Delta f_T^2 \cdot N_p^2} \tag{2.7}$$

The next step consisted in calculating the distance b_x between the centred real source position $[S_c]$ and the calculated side position $[S_A]$:

$$b_x = r_x - e_x \tag{2.8}$$

$$\Delta b_x = \sqrt{\Delta r_x^2 + \Delta e_x^2} \tag{2.9}$$

The error Δr_x was assumed to be $\pm 0.05m$.

The length b_x could then be used to calculate the angle α_h:

$$\alpha_h = \arctan \frac{b_x}{a} \tag{2.10}$$

and its error:

$$\Delta\alpha_h = \sqrt{(\frac{a}{a^2 + b_x^2})^2 \cdot \Delta b_x^2 + (\frac{b_x}{a^2 + b_x^2})^2 \cdot \Delta a^2} \tag{2.11}$$

Finally θ_h, being the horizontal angle from the real centre position of the source S_c to the real side position S_L was calculated with:

$$\theta_h = \arctan \frac{r_x}{a} \tag{2.12}$$

and its error:

$$\Delta\theta_h = \sqrt{(\frac{a}{a^2 + r_x^2})^2 \cdot \Delta r_x^2 + (\frac{r_x}{a^2 + r_x^2})^2 \cdot \Delta a^2} \tag{2.13}$$

Finally all the variables to calculate φ_h and $\Delta\varphi_h$ were given.

$$\boxed{\varphi_h = \theta_h - \alpha_h} \tag{2.14}$$

$$\boxed{\Delta\varphi_h = \sqrt{\Delta\theta_h^2 + \Delta\alpha_h^2}} \tag{2.15}$$

An example of the final graphical results for the horizontal deviation angle can be observed in figure 2.12. The dashed lines indicate the $\pm 3°$ interval.

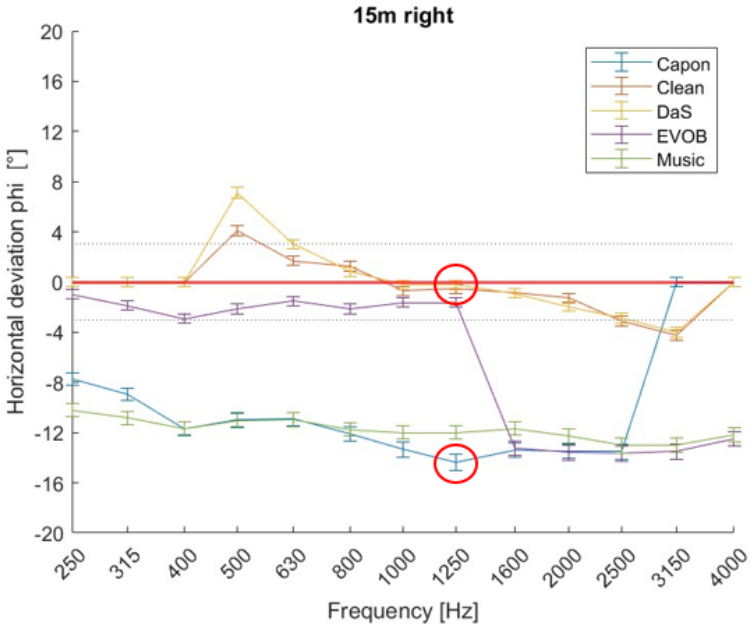


Figure 2.12: Horizontal deviation angle at 15m distance to the source. The source was placed to the right. The two circles mark a good and a bad result. The associated images can be seen in figure 2.15.

Vertical deviation angle φ_v

Image 2.13 shows the underlying concept for the calculation of the vertical deviation angle φ_v between the real source side position $S_L(left)$ and the one detected by the different algorithms S_A.

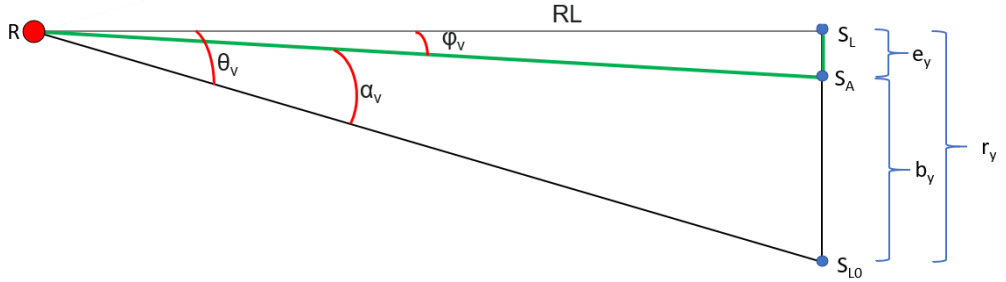


Figure 2.13: Scheme for the vertical deviation angle φ_v calculation.

RL is the distance from the receiver R and the source position $S_L(left)$. For the right position S_R RL has the same size, whereas for the centre position S_C

$RL = a$. The receiver R and the side position $S_L(left)$ where placed at the same height of 1.2 m from the ground. The difference from the calculated source position S_A and the real source position S_L is e_y in pixel. It was transformed into meters with the same calculation as e_x in the description of the horizontal deviation angle (2.6 and its error 2.7); r_y is the distance to the ground and is equivalent to 1.2 m; S_{L0} marks the point on the ground. Knowing the distance from the source a (remember the scheme for the horizontal deviation in figure 2.11) and the horizontal angle θ_h it was possible to calculate the distance RL with:

$$RL = \frac{a}{cos(\theta_h)} \tag{2.16}$$

Based on the error propagation equations in [38] the associated error ΔRL was calculated as follows:

$$\Delta RL = \sqrt{\frac{\Delta a^2}{cos^2(\theta_h)} + \frac{a^2}{cos^4(\theta_h)} \cdot \Delta cos^2(\theta_h)} \tag{2.17}$$

With the values of RL and the difference e_y between the calculated height position of the source S_A and the real height position of the source S_L it was possible to calculate the vertical deviation angle φ_v for all distances and source positions:

$$\varphi_v = arctan\frac{e_y}{RL} \tag{2.18}$$

and its error

$$\Delta\varphi_v = \sqrt{(\frac{RL}{e_y^2 + RL^2})^2 \cdot \Delta e_y^2 + (\frac{e_y}{e_y^2 + RL^2})^2 \cdot \Delta RL^2} \tag{2.19}$$

An example of the final graphical results for the vertical deviation angle can be observed in figure 2.14. The dashed lines indicate the $\pm 3°$ interval.

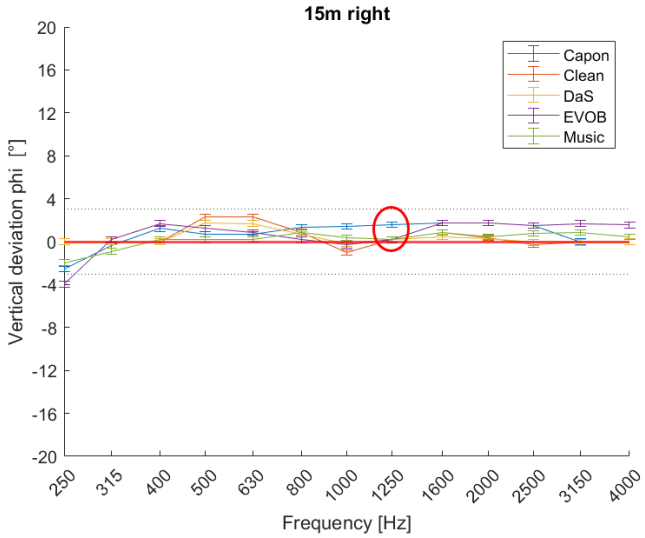
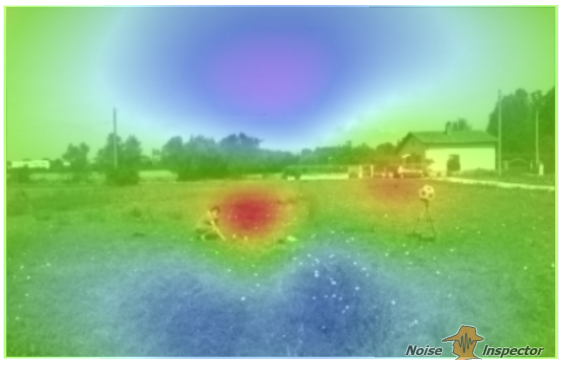
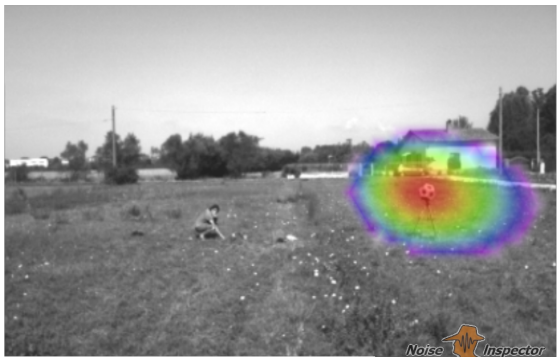


Figure 2.14: Vertical deviation angle at 15m distance to the source. The source was placed to the right. The circle marks the result for the *Capon* and *DaS* algorithm at the frequency of 1250Hz. The vertical deviation angle is not as wide as the horizontal angle. The associated images can be seen in figure 2.15.



(a)



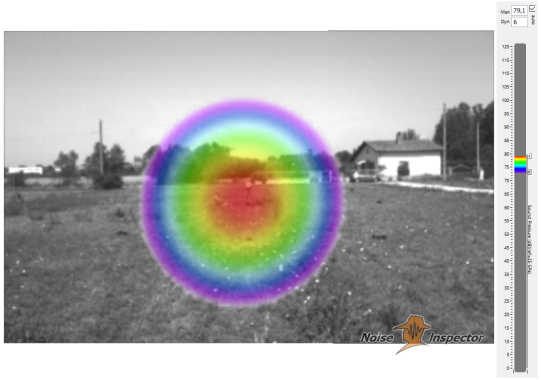
(b)

Figure 2.15: Two images created with the acoustic camera for the source at 15 m distance displaced to the right. Image a: Evaluation with the capon algorithm at the third octave frequency band of 1250 Hz, a poor result. Image b: Evaluation with the DaS algorithm at the third octave frequency band of 1250 Hz, an accurate result.

2.7 Area of maximum Sound Pressure Level [SPL]

The software *Noise Inspector* creates an image with a color spectrum where each color represents one sound pressure level. The results are similar to the well known images from thermal cameras.

(a)



(b)

Figure 2.16: Two images created with the acoustic camera for the source at 30 m distance in the centre position. Image a: Evaluation with the DaS algorithm at the third octave frequency band of 315 Hz, a valuless result. Image b: Evaluation with the DaS algorithm at the third octave frequency band of 800 Hz, an acceptable result.

In our case, the maximum sound pressure level detected by the acoustic camera is displayed with red colour. The colour scale has 6 colours (red, yellow, green, light blue, dark blue and purple) and a ΔSPL of 6 dB. Hence, each color represents $1dB$. An example of a partly satisfactory result can be observed in picture 2.16b. It was produced by the DaS algorithm at 800 Hz with the source at a height of 1.2 m and 30 m from the antenna. In comparison an image of an unsatisfactory result is shown in picture 2.16a. It is the result of an analysis for

the same source position and algorithm but evaluated at 315 Hz.

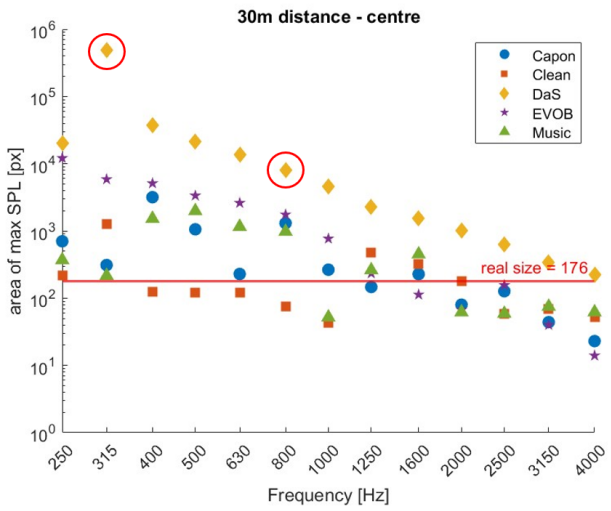


Figure 2.17: Graph with the different calculated sizes of the maximum SPL area for the centre source position 30m from the acoustic camera. The marked points are the result of the *DaS* algorithm at 315 Hz and 800 Hz.

For the previous analysis of the deviation angle we only considered the exact centre of the detected source. However, another relevant parameter for comparing the algorithms is the size of the calculated, maximum SPL area. If the whole image is red the accuracy could seem elevated, because the centre was located, but in fact the user does not get a useful information about the exact position of the source in the image. Hence, the size of the area of the maximum SPL was analysed. For this purpose, the active bins of the previously calculated mask were summed up and compared to the number of bins of the real source. Of course, the distance to the source and thus the changing size of the dodecahedron in the image was taken into account and included in the evaluation of the accuracy of the algorithms. The results were 12 representative images; one for each distance and source position. One example is shown in figure 2.17. The red circled points in the graph mark the size of the area of maximum SPL calculated for the image created with the *DaS* algorithm at 315 Hz and 800 Hz, namely figure 2.16.

3 Results and Discussion

3.1 Success rate

The first step was to calculate the success rate for each algorithm. This calculation showed very positive results. *DaS* and *Music* algorithm showed a success rate of 100% for every distance. The least efficiency is given by the *Evob* algorithm.

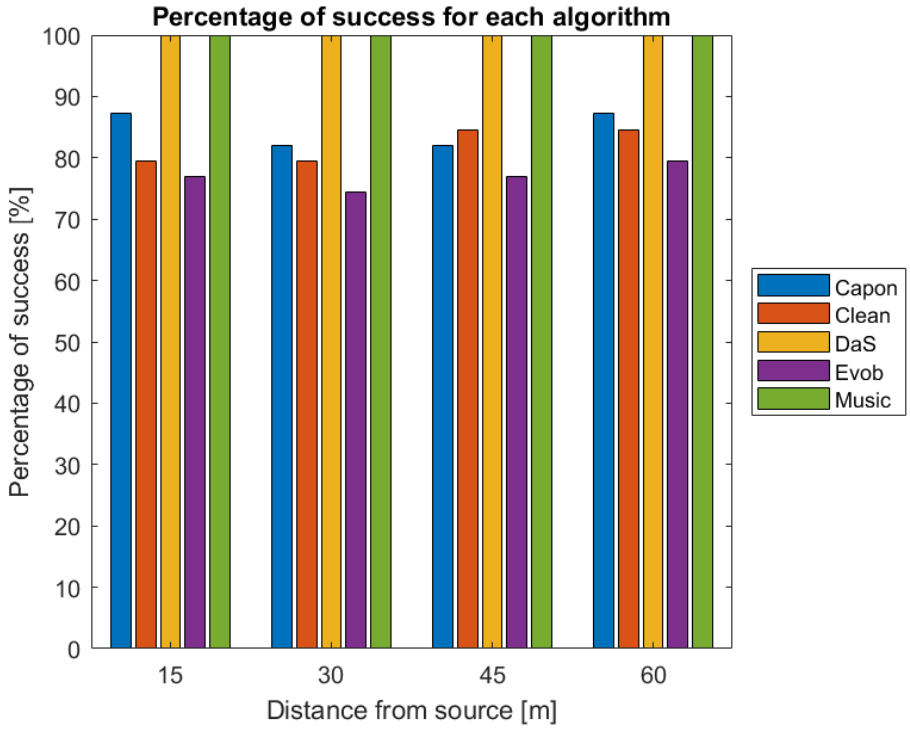


Figure 3.1: Percentage of success for all the distances.

3.2 Bland Altman diagrams

As mentioned before, each algorithm is represented by a different colour (see table 2.4).
For the measurement at 15m distance very good results can be observed when the source is in the middle. Only a few pixels of difference are between the groundtruth centre and the calculated one. Only the results calculated by the algorithm *Clean SC* and *DaS* show data lying outside the limits.
At a distance of 30m and in centered position the difference is close to zero. Off-axis to the left and right the difference is considerable for the x-axis values. However it is very small for the y-axis values. Outside the limits lay again only results from the computation with the algorithms *Clean SC* and *DaS* and in one case also *Evob*.
At the distance of 45m the largest difference is detected in the data set where the source is displaced to the right. Again *Clean SC*, *DaS* and *Evob* have some data lying outside the limits.
At the distance of 60m the results are very similar to the ones at the other distances. The centered position of the source shows the most accurate results for x- and y-axis. Left and right displacements of the source lead to more inaccurate results. At this distance all algorithms held at least one point outside the limits.

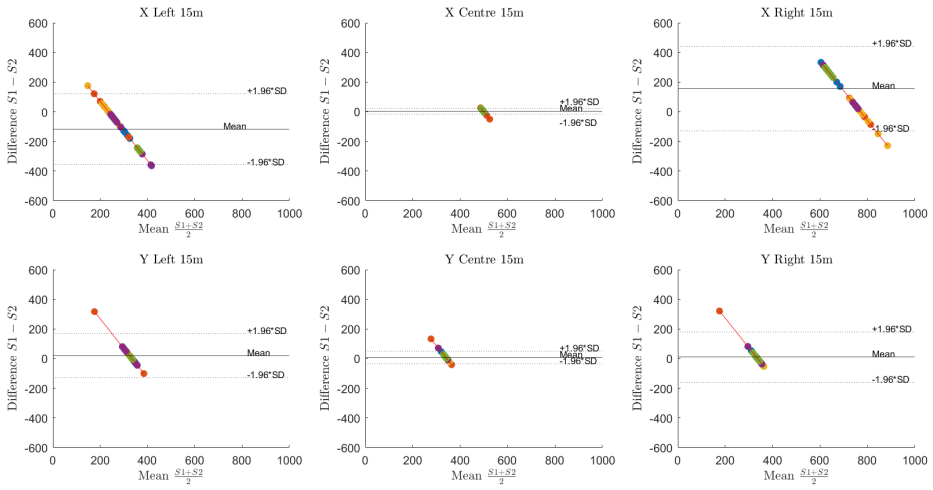


Figure 3.2: Bland Altman diagrams for 15 m distance to the source.

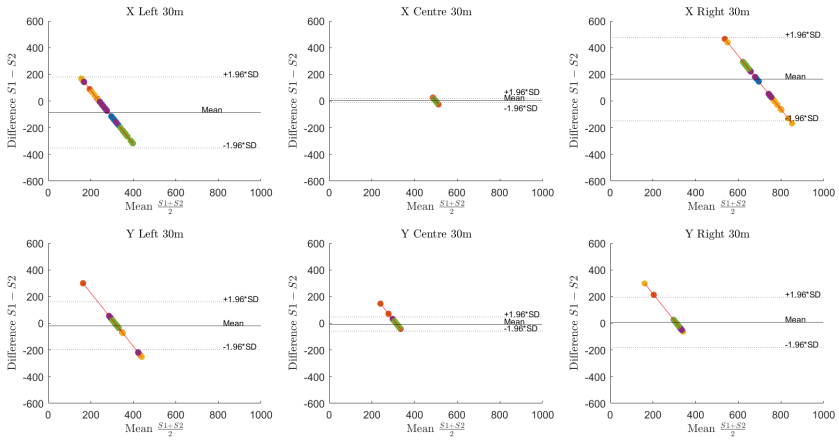


Figure 3.3: Bland Altman diagrams for 30 m distance to the source.

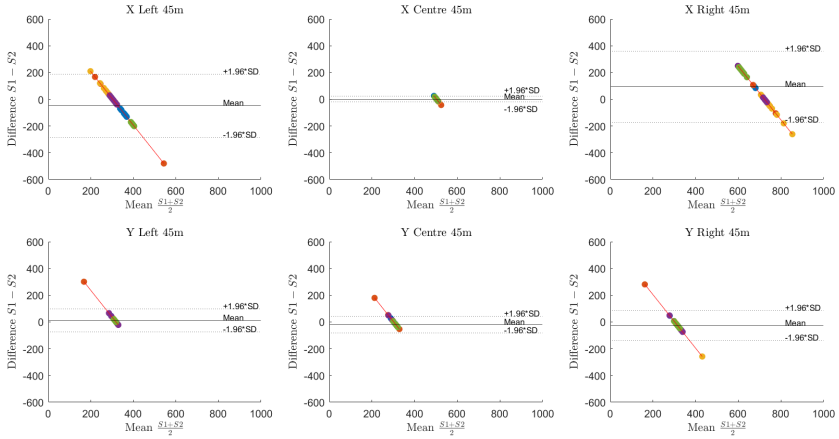


Figure 3.4: Bland Altman diagrams for 45 m distance to the source.

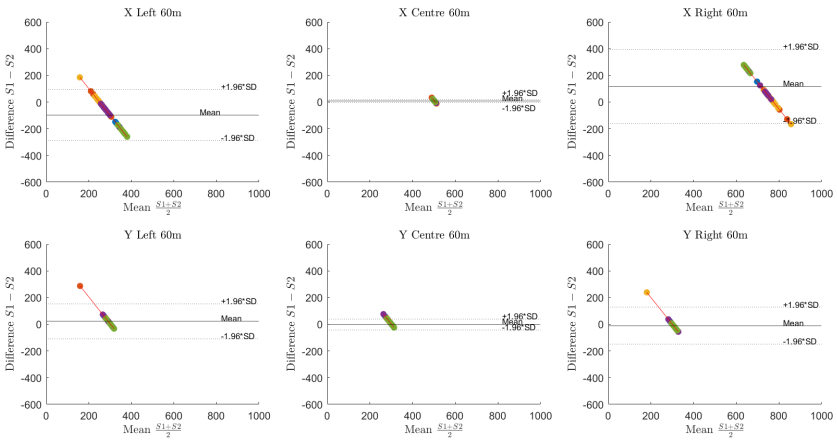


Figure 3.5: Bland Altman diagrams for 60 m distance to the source.

3.3 Deviation angle

After transforming the distance r_x between the central source position and the lateral source positions in the image from pixel to meter and then calculating the horizontal angle Θ_h between these two points (see image 2.11), it was found that the angle was slightly different from the planned angle $\Theta_p = 10,5^{\circ}$. Table 3.1 reports the angle for each distance. These values have been used for all further calculations of the horizontal and vertical deviation angle.

Table 3.1: Angle Θ_{ac} between the central position of the source and the lateral source positions seen from the acoustic camera. See also image 2.11.

distance a [m]	**angle θ_{ac} [°]**
15	12,77
30	12,00
45	11,67
60	10,50

The dotted lines in all the reported deviation angle graphs in this section mark the $\pm 3°$ interval for the results.

3.3.1 Horizontal deviation angle

The sign of the horizontal deviation angle gives information about whereas the calculated source position S_A was placed to the left or to the right in relation to the real source position $S_L(left), S_R(right)$ or $S_C(centre)$. The negative sign indicates that S_A lays further to the left than the real source position.

Capon algorithm

The results for the *Capon* algorithm are very similar at each distance. For the source at the centre position the accuracy is high, the deviation angle never reaches the value of 1°.
Moving the source to the right results in a bigger horizontal deviation angle. It fluctuates between $-8°$ and $-15°$ at every distance with minima of $-3°$ for some frequency bands.
For the left direction the results are slightly better. The horizontal deviation angle is between 3° and 8°.

Clean SC algorithm

Also the *Clean SC* algorithm is accurate for the central positions of the source. Only at the distance of 15m for the two lowest frequency bands the deviation angle raises to 2°.
For the left position of the source a misalignment from $-3°$ to $-7°$occurs at the frequency of 500 Hz and then narrows towards the real source position again for higher frequencies.
To the right side a similar mismatch can be observed from 500 Hz onward. A striking value can be observed at the distance of 30m for the frequency band of 3150 Hz: $-18°$. This is the highest value for this algorithm.

Delay and Sum (DaS) algorithm

The *DaS* algorithm generated very similar results to the clean algorithm especially for the mid third octave bands.
It also has noteworthy low value of $-7°$ for the left side position of the source and the 500 Hz frequency band. Almost the same value but with a positive sign is calculated for the same source position at 2500 Hz. For all the other analysed source positions the 500 Hz band produces the worst results.
The values follow the same tendency for both the side deviations and all the distances. The results for the centred source position are the most accurate, never exceeding a 1° horizontal deviation.

Evob algorithm

The results for the central position of the source are accurate also for this algorithm. They never reach the 1°.
However, there are some inaccuracies of around $\pm 2°$-$\pm 3°$ for the side positions at lower frequencies already at the closest distance of 15 m and running through the other distances.
The most noticeably errors occur at higher frequencies. For the right side already from 1600 Hz, where the deviation angle lays from $-13°$ to $-9°$. For the left side from 2500 Hz where the deviation angle reaches 17°.

Music algorithm

This algorithm shows the least fluctuating lines in the graph, but also the lowest accuracy.
While the deviation angle for the centre position is very low, the left and right deviation angle is quite high at all distances. On the left side for 15 m and 30 m distance the angles lay between 10° and 15° and for 45 m and 60 m between 7° and 9°. On the right side at 15 m and 30 m curve lays between $-10°$ and $-12°$ and at 45 m and 60 m between $-10°$ and $-7°$.
This algorithm provides the worst information about the real position of the source.

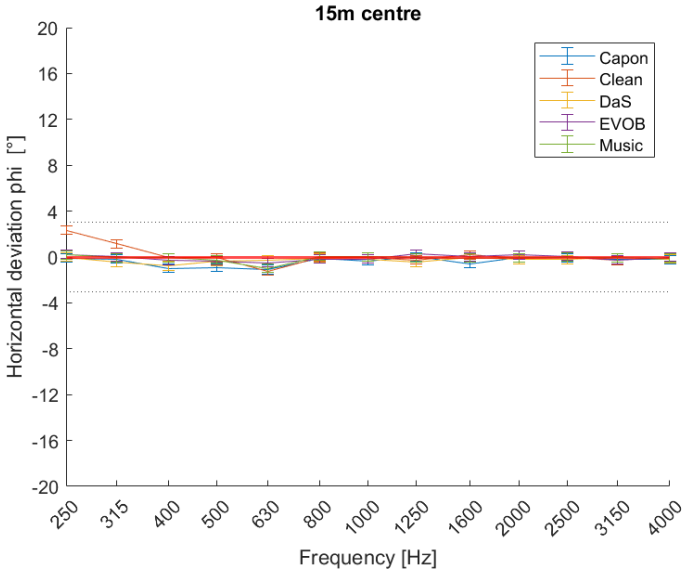


Figure 3.6: Horizontal deviation angle at 15m distance to the source. The source was placed in the centre.

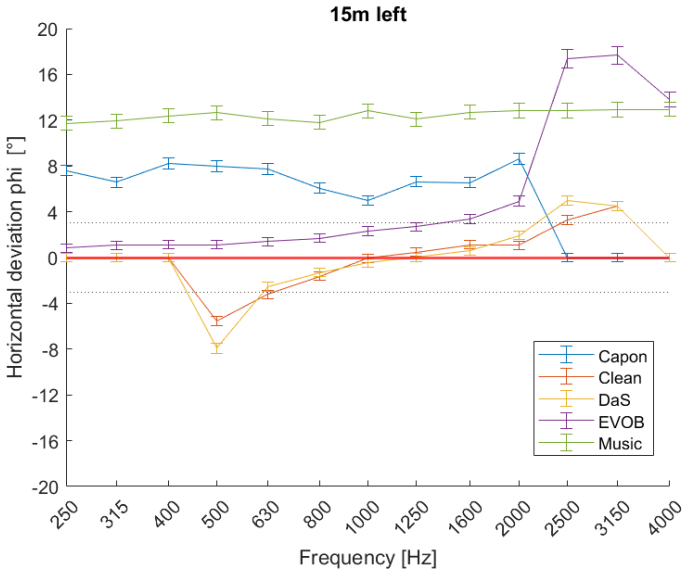


Figure 3.7: Horizontal deviation angle at 15m distance to the source. The source was placed to the left.

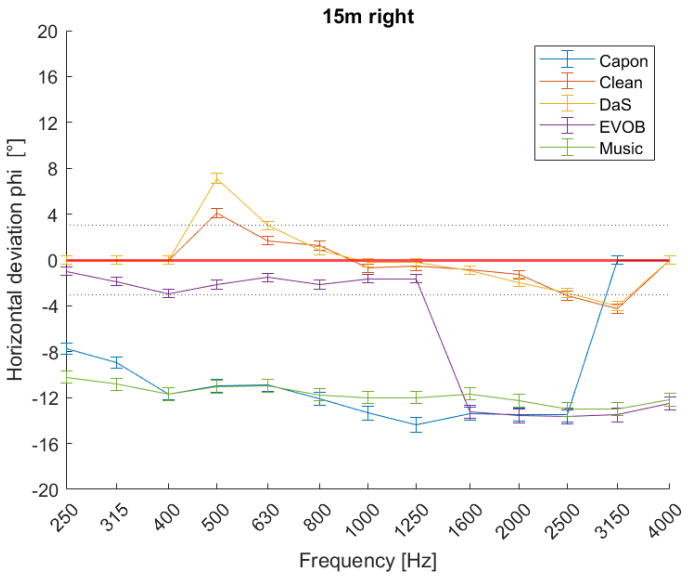


Figure 3.8: Horizontal deviation angle at 15m distance to the source. The source was placed to the right.

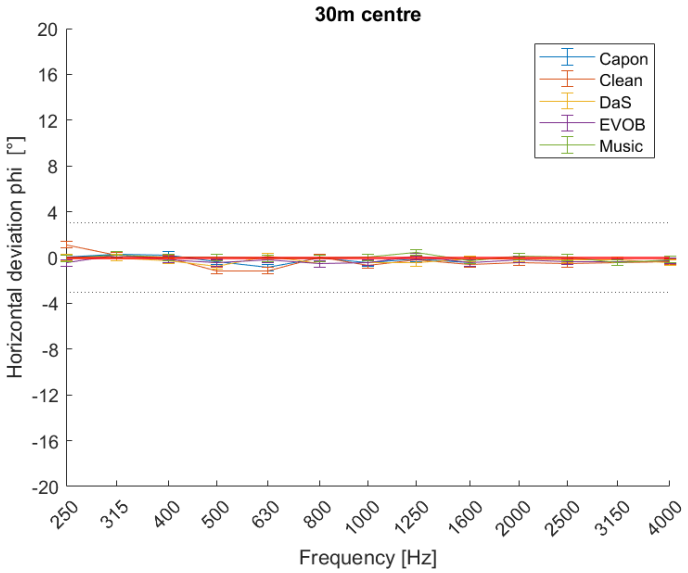


Figure 3.9: Horizontal deviation angle at 30m distance to the source. The source was placed in the centre.

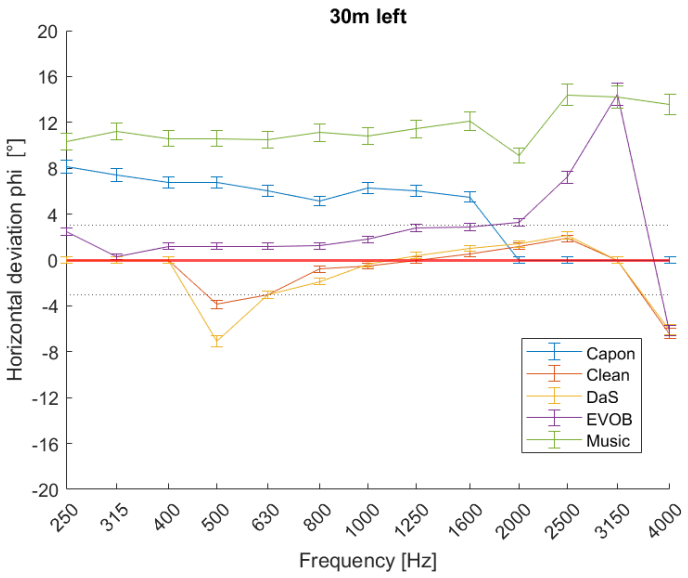


Figure 3.10: Horizontal deviation angle at 30m distance to the source. The source was placed to the left.

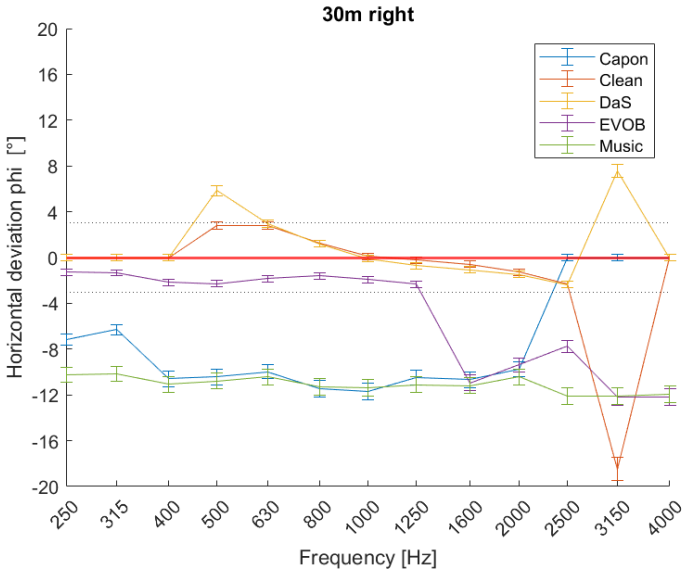


Figure 3.11: Horizontal deviation angle at 30m distance to the source. The source was placed to the right.

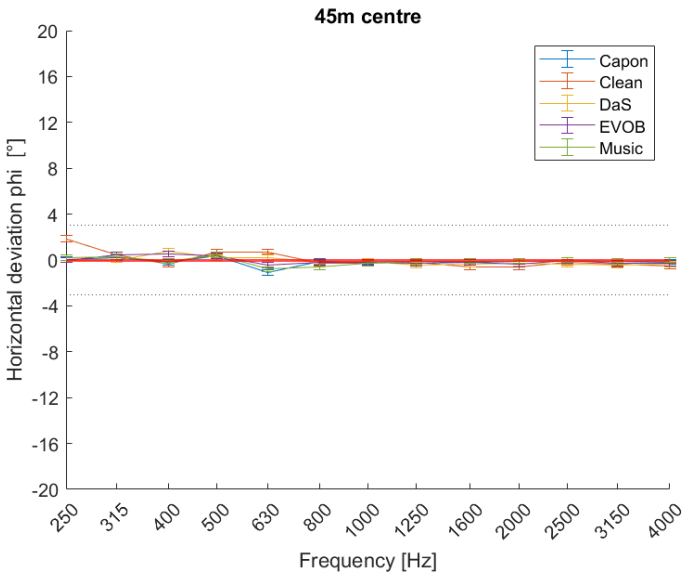


Figure 3.12: Horizontal deviation angle at 45m distance to the source. The source was placed in the centre.

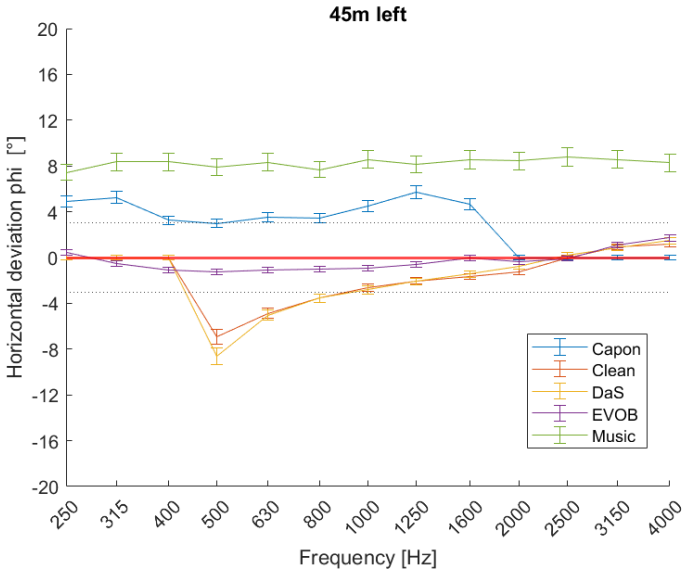


Figure 3.13: Horizontal deviation angle at 45m distance to the source. The source was placed to the left.

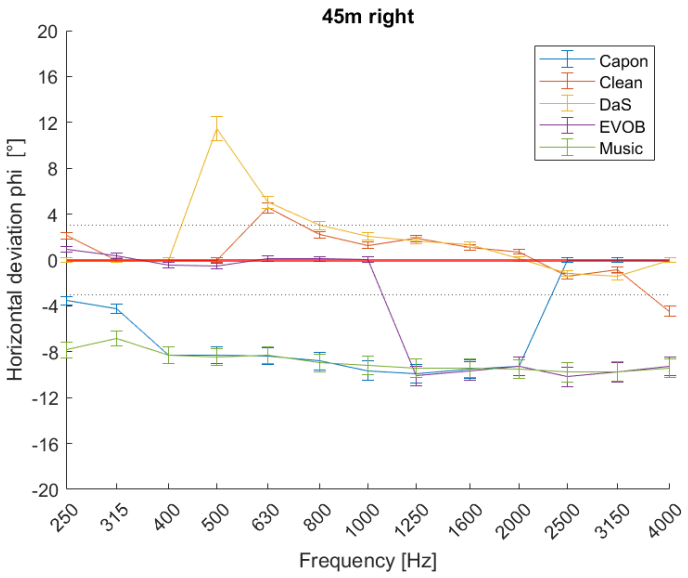


Figure 3.14: Horizontal deviation angle at 45m distance to the source. The source was placed to the right.

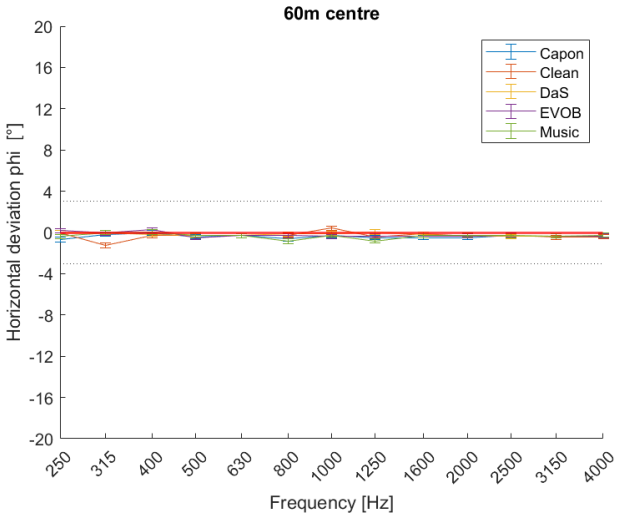


Figure 3.15: Horizontal deviation angle at 60m distance to the source. The source was placed in the centre.

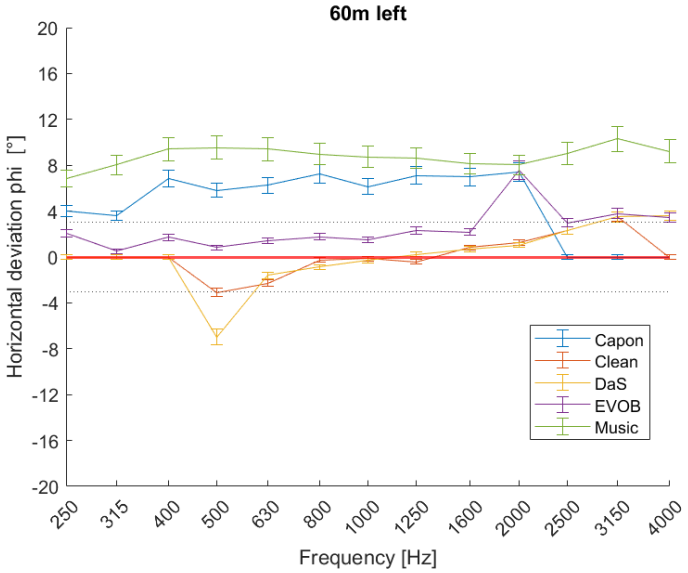


Figure 3.16: Horizontal deviation angle at 60m distance to the source. The source was placed to the left.

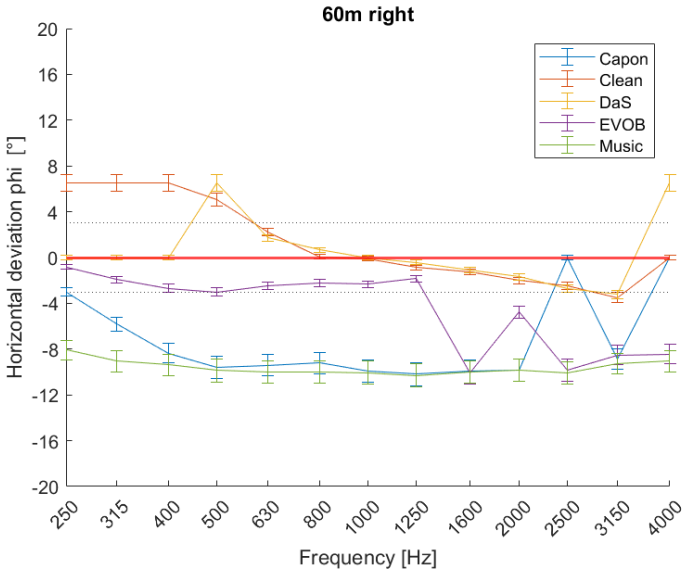


Figure 3.17: Horizontal deviation angle at 60m distance to the source. The source was placed to the right.

3.3.2 Vertical deviation angle

The sign of the vertical deviation angle gives information about whereas the calculated source position S_A was placed higher or lower than the real source position $S_L(left), S_R(right)$ or $S_C(centre)$. The negative sign indicates that S_A lays further up than the real source position.

Capon algorithm

Generally speaking the vertical deviation error is much smaller than the horizontal one and the curves are less fluctuating. This is also the case for the *Capon* algorithm. It follows the same tendency over all the distances and orientations (left, centre, right): it starts at around $-2°$ at 250 Hz and then moves closer to 0° and lays around 2° from 400 Hz onward. For the central position the accuracy is even higher, laying constantly around 1°-1,5°.

Clean SC algorithm

The *Clean SC* algorithm produces the worst results. Nonetheless, almost all the values are close to the 0°-line.
Some exceptions are seen in the centre at 15 m in the 250 Hz frequency band ($-6°$), to the left side at 15 m at 2500 Hz (5°), in the centre at 30 m in the 250 Hz frequency band ($-7°$), to the left side at 30 m in the 4000 Hz frequency band (10°), to the right side at 30 m in the 3150 Hz frequency band ($-10°$) at 45 m in the centre at 250 Hz ($-8°$), at 45 m to the right at 250 Hz ($-12°$) and at 60 m to the right for 250 Hz, 315 Hz and 400 Hz ($-9°$).

Delay and Sum (DaS) algorithm

The *DaS* algorithm follows the curve tendency of the *Clean SC* algorithm without the anomalies for distinct frequency bands. Only two exceptional values can be observed for the distance of 30 m to the left at 4000 Hz with 10° vertical deviation and for the 60 m to the right at the same frequency band with $-9°$ vertical deviation.

Evob algorithm

The *Evob* algorithm fluctuates between $-3°$ and $+3°$. Tendentially the values lay around $-3°$ for the lowest frequency band and then increase towards $0°$ or little above for the mid to high frequencies. Two exceptions are given for the 15 m distance to the left side at 2500 Hz where the deviation angle is $-3°$ and at 3150 Hz where the deviation angle is $-2°$. Also the further distance of 30 m towards the left side has an exceptional value of $10°$ at 4000 Hz.

Music algorithm

The *Music* algorithm shows more valuable results than for the horizontal deviation angle. There are no significant anomalies. All the values lay constantly around $\pm 1°$ for all distances and all orientations. When observing only the vertical deviation angle, it could seem that the *Music* algorithm is the most accurate. However, we have seen the very poor results it produces for the horizontal deviation angle.

Conclusion

The results for the horizontal and vertical deviation angles can be summed up as followed:

- The vertical deviation angle is significantly smaller than the horizontal deviation angle.
- For the vertical deviation angle there is no remarkable difference between the results at the centre position or the results from the side positions.
- On the other hand, for the horizontal deviation angle the difference between central position and side positions is noteworthy.
- There is no noticeable increase in the deviation angles for the increasing distances.

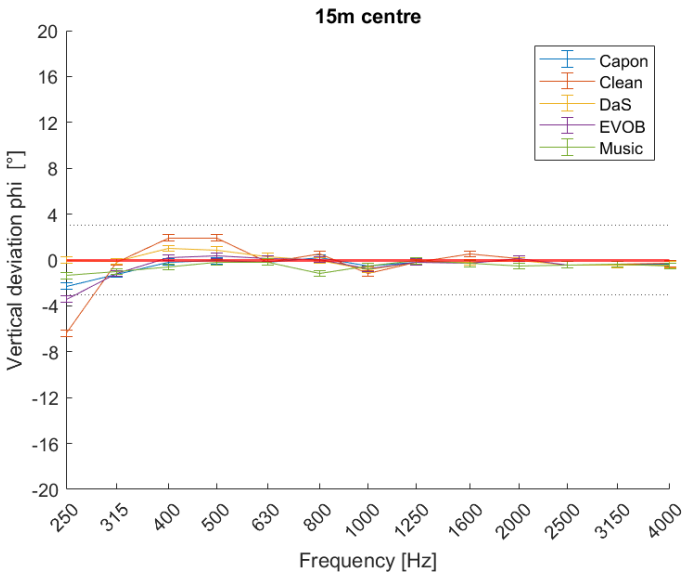


Figure 3.18: Vertical deviation angle at 15m distance to the source. The source was placed in the centre.

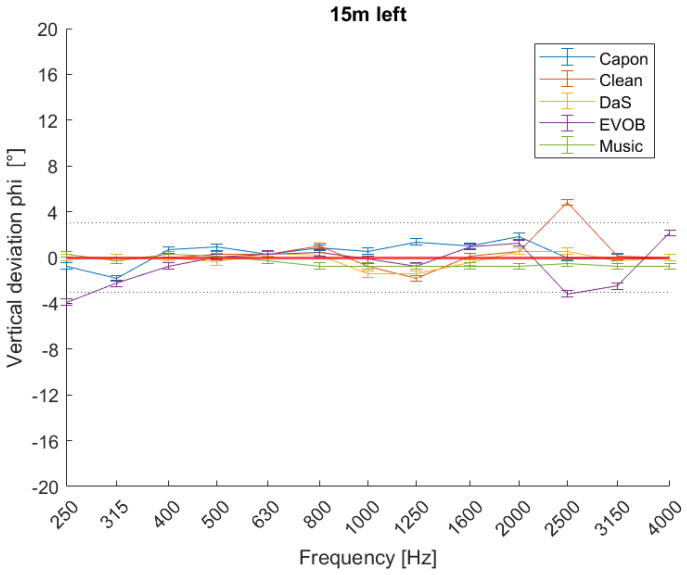


Figure 3.19: Vertical deviation angle at 15m distance to the source. The source was placed to the left.

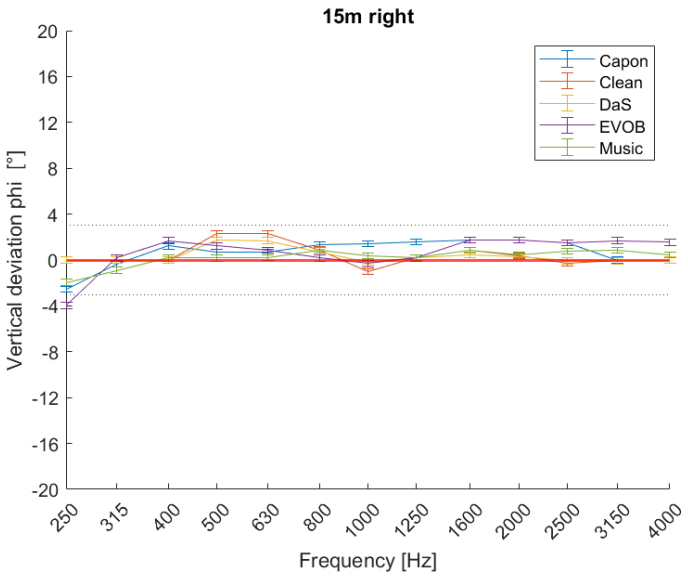


Figure 3.20: Vertical deviation angle at 15m distance to the source. The source was placed to the right.

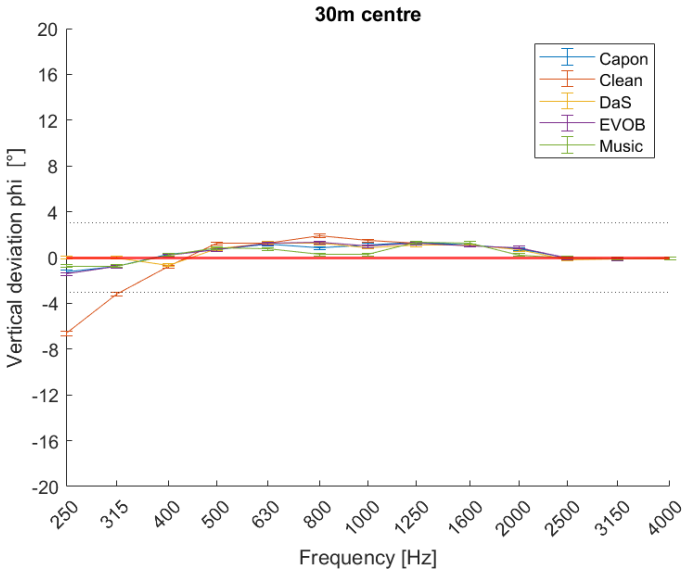


Figure 3.21: Vertical deviation angle at 30m distance to the source. The source was placed in the centre.

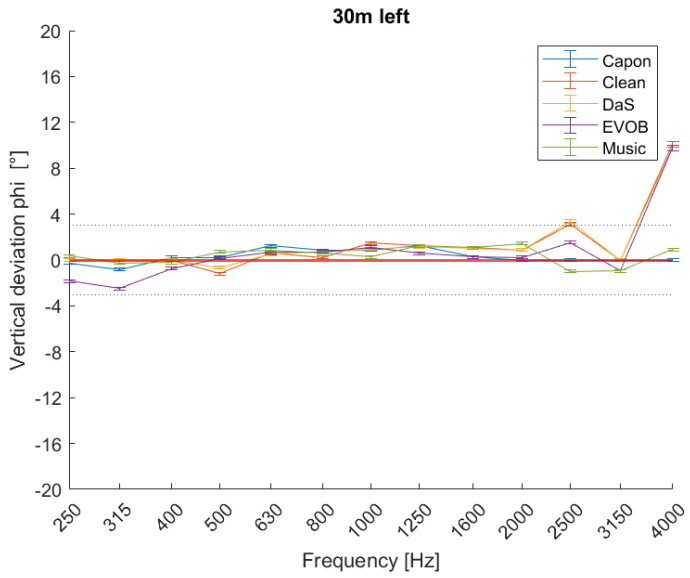


Figure 3.22: Vertical deviation angle at 30m distance to the source. The source was placed to the left.

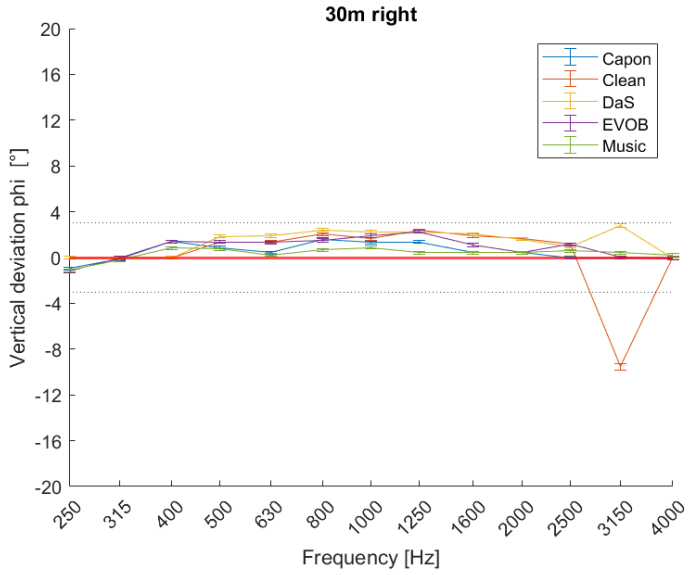


Figure 3.23: Vertical deviation angle at 30m distance to the source. The source was placed to the right.

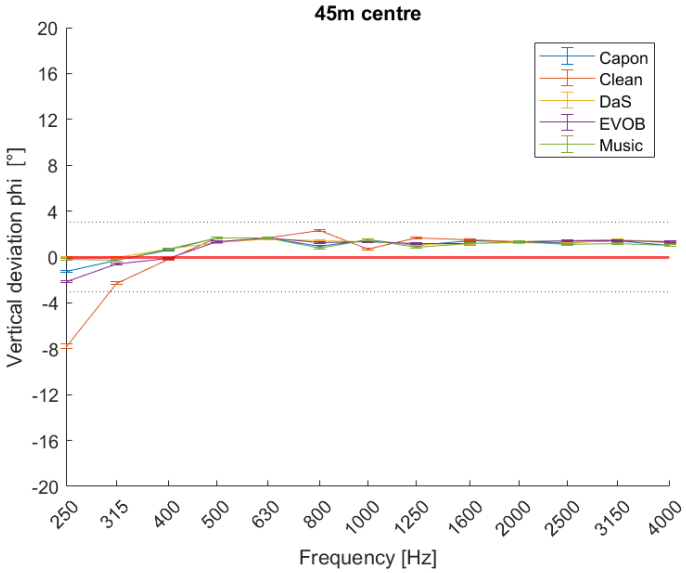


Figure 3.24: Vertical deviation angle at 45m distance to the source. The source was placed in the centre.

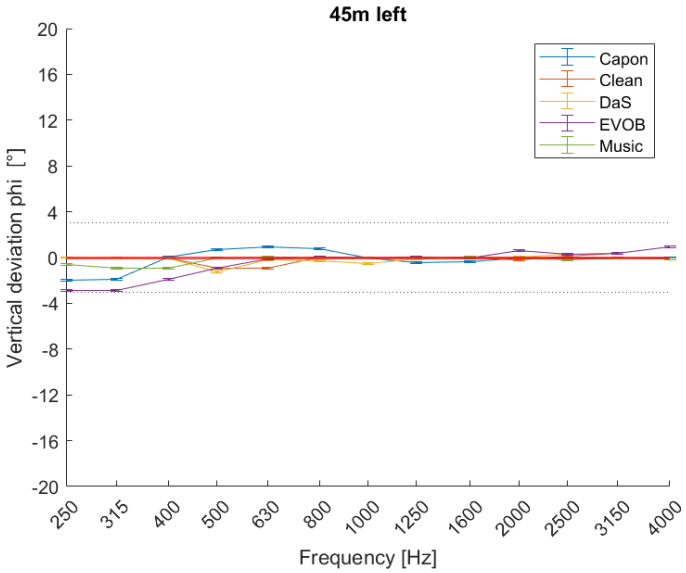


Figure 3.25: Vertical deviation angle at 45m distance to the source. The source was placed to the left.

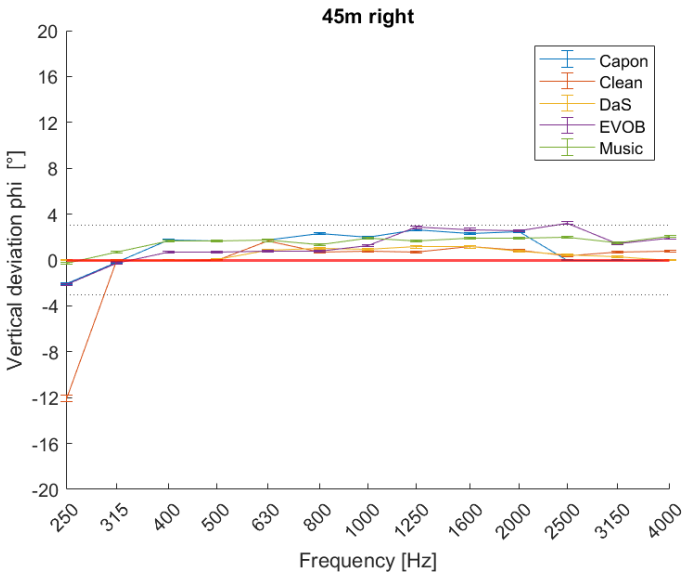


Figure 3.26: Vertical deviation angle at 45m distance to the source. The source was placed to the right.

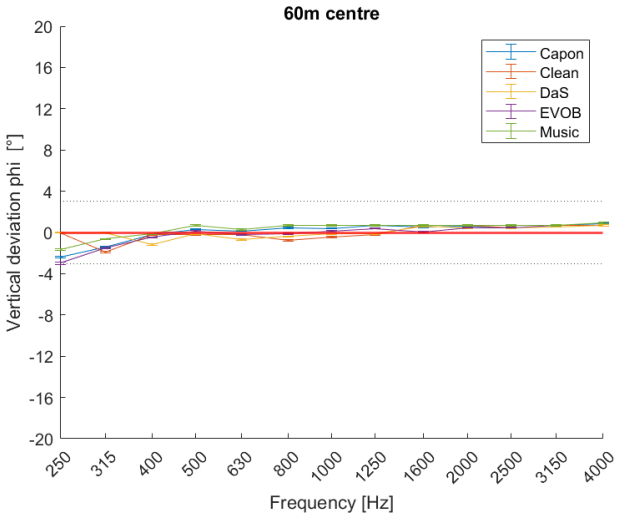


Figure 3.27: Vertical deviation angle at 60m distance to the source. The source was placed in the centre.

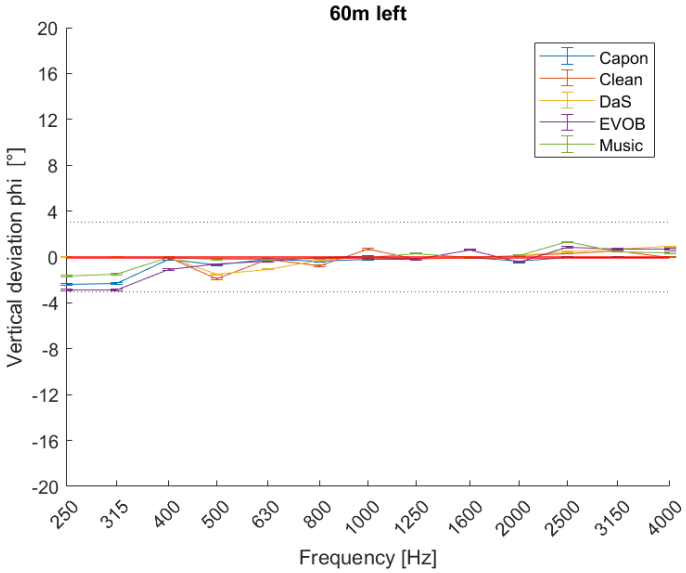


Figure 3.28: Vertical deviation angle at 60m distance to the source. The source was placed to the left.

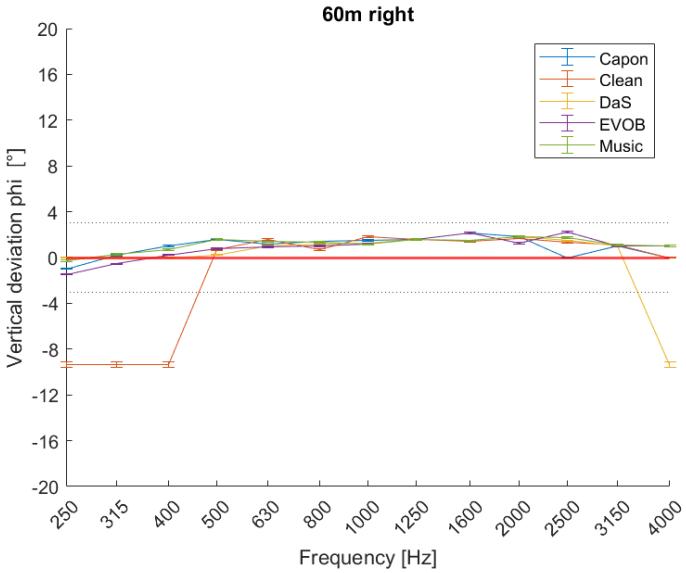


Figure 3.29: Vertical deviation angle at 60m distance to the source. The source was placed to the right.

3.4 Area of maximal Sound Pressure Level

As described in section 2.7 from page 44, the area of maximal sound pressure level was calculated to determine the efficiency of each algorithm type. The results of the single algorithms were joined in a graph depending on their distance to the source and variance from the source. Hence, for this analysis 12 graphs were generated. The area of maximal SPL in pixel was plotted over the third-octave frequency bands from 250 Hz to 4000 Hz. For a better legibility the area of the real source was drawn as a red constant line over the entire frequency spectrum. For a simplification of the calculation we assumed a constant radiation pattern of the dodecahedron-shaped source over the whole frequency range.
The most noticeable fact is that, even though the real size of the source diminishes with the increase of the distance between source and acoustic camera, the calculated area of maximal SPL is not directly affected by this decrease. This suggests a lack of reliability with regard to the recognition of the distance from the source to the acoustic camera. Nonetheless, all the 12 graphs are described in the following paragraphs.

Capon algorithm

The *Capon* algorithm does not show remarkable results at the distance of 15m with the source in the centre and left position. Still, it creates acceptable results for the right shift. Also, it should be mentioned, that the algorithm underestimates the SPL area for the centre position. So, in combination with an accurate location detection, the results can be considered as valuable.
The translocation to a distance of 30m from the acoustic camera yields more accurate results for the centre position throughout most of the observed frequency range. However, the left and right shift of the source lead to an overestimation of the maximum SPL area.
Further distance increase to 45m especially changed the result for the centre position. The calculated area increased in comparison to the real size. The left and right shift showed very similar results to the previous distance setup, with the respective ratio due to the smaller real size in the picture.
The results in the graph for the 60m distance to the acoustic camera are consistent with the previous trend for the 45m distance.
In summary, the *Capon* algorithm overestimates the size of the source and tends to work better when the source is centrally targeted.

Clean SC algorithm

The *Clean SC* algorithm shows an evident underestimation of the max SPL area for the centre position (15 m). All values lay notably below the real source size. The shift to the left side gives results, that are closer to the real size of the source in the 1000Hz, 1250Hz, 1600Hz and 2000Hz frequency bands. The outermost values of the frequency range show a very big value for the calculated size of the source. Similarly, the shift to the right side produces poor results for the low frequencies up to 400Hz and the 4000Hz frequency band, some acceptable results for the frequencies from 500Hz to 1000Hz and some good results for the 1250Hz, 1600Hz and 2000Hz frequency band.
Satisfyingly, the central position of the source at a distance of 30 m provides sufficient results in most of the analysed frequency bands. On the other hand, the left and right variation of the source position results in strongly fluctuating data points.
Next, the distance of 45m leads again to better results for all three different displacement positions. The central position shows good results especially in the frequency bands from 800Hz to 3150Hz, with a slightly higher overestimation for the lower frequencies. Also the left displacement of the source creates good results almost throughout the whole frequency range, with the exception of the 250Hz, 400Hz, 500Hz and 630Hz frequency bands. The dislocation to the right produces very accurate results in the 315Hz, 400Hz, 500Hz and 2500Hz frequency bands and acceptable results in the remaining frequency bands with an exception at 250Hz, 630Hz and 800Hz.
The furthest distance of 60m from the acoustic camera yields fair results in the lower frequencies up to 1000Hz and again at 2000Hz, 3150Hz and 4000Hz. The results in the intermediate frequency bands slightly overstate the size of the source. Again, the data points of the left and right shift are fluctuating much more than the central position. However, for both side orientations the results in the 500Hz, 630Hz, 800Hz, 1000Hz, 2500Hz and 3150Hz are valuable.
In conclusion, the *Clean SC* algorithm underestimated the size of the source especially in the closer setups at a distance of 15m and 30m. This underestimation, in conjunction with the calculated location of the source, is more valuable than an overestimation and thus an inaccurate location point in the picture. Considering the four distances and three alignments of the source, the *Clean SC* algorithm produced the lowest values for the max SPL area compared to all other algorithms.

Delay and Sum [DaS] algorithm

Although this algorithm gave the most consistent results, it is not a satisfactory outcome: consistent trends are observed across various distances for each alignment. The calculated area of max SPL for the central alignment follows a linear decrease from low to high frequencies. In contrast, the left and right alignments exhibit a linearly increasing trend from 250Hz to 500Hz, followed by a linear decrease towards higher frequencies. Furthermore, the actual size of the source is overestimated in nearly all data points.
Conclusively, the *DaS* algorithm does not give exact results. However, the tendency remains the same for all changes in distance and alignment, so it is possible to predict and calculate the error.

EVOB algorithm

The results of the closest (15m) and central position of the source are close to the aim size at 800Hz with a further decrease towards the higher frequencies. Changing the alignment towards the left side does not lead to major changes. However, altering the position to the right side causes an increased value of the deviation angle especially at higher frequencies.
At a distance of 30m the linearity is clearly visible for the central position. The results from 1250Hz to 2500Hz are valuable. The trend for the left side offset is similar, however, the acceptable results begin from the 1600Hz frequency band. The right side shift also produces good results from the 1600Hz frequency band onward, but the decrease is steeper as for the left side.
A very similar graph shows the results for the 45m distance and central orientation: a linear decrease towards the higher frequencies and the real value line is reached at 2000Hz. The left side alignment only reaches the real size line at 4000Hz whereas the right side alignment shows acceptable results beginning from 2000Hz.
The furthest distance (60m) produces once again the poorest results. The ones for the central position are sufficient at 2500Hz, 3150Hz and 4000Hz, while for the left and right position of the source the results do not reach the value of the real size. For both directions the results at 3150Hz and 4000Hz can be considered sufficiently accurate.
To conclude, the *EVOB* algorithm generates similar results to the *DaS* algorithm when speaking of tendency. However, the distance from the imaginary, decreasing straight line along the data points and the constant real size line is always smaller for the *EVOB* algorithm. Therefore, it can be said that the algorithm has a higher accuracy.

Music algorithm

The nearest (15m) and central position did produce satisfactory results. Albeit, most results lay below the constant real size line. In comparison, the left side results yield positive outcomes near the optimal value for the low frequencies from 250Hz to 630Hz and obtaining lower results for the remaining frequency bands. For the right side shift the values at 630Hz, 800Hz, 1600Hz, 2000Hz and 2500Hz are exemplary.
Moving to a 30m distance creates inconsistent results for the central position of the source. One good result can be observed at 315Hz, again at 1250Hz and some underestimated values for 1000Hz and from 2000Hz upwards. In contrast, the left and right shift follow approximately constant decrease towards higher frequencies without ever crossing the real size magnitude.
For the successive distance of 45m the central and right alignment show similar hypothetical curve progressions. Disappointingly, the right side offset produces unremarkable results. Although the left-side orientation yields remarkable outcomes for the 800Hz, 1000Hz, 1250Hz, and 1600Hz frequency bands, it tends to overestimate the magnitude for the remaining frequency bands.
Once again, the central position attains the best results for the 60m distance in comparison to the left and right shift; however, the outcomes still fall short of satisfactory levels. Only the values at 1000Hz, 2000Hz, 3150Hz and 4000Hz achieve the real size constant. Nonetheless, the left and right shift produce inadequate outcomes and fail to approach the true value of the size. In conclusion, it is important to outline that the results were noteworthy only for the distances of 15m and 30m.

Conclusion

In summary, it can be said that the *Clean SC* algorithm performed the best, consistently producing small areas of maximum sound pressure level (SPL). The *Capon* algorithm showed high fluctuation within this data sets. The inconsistency made it difficult to rely on the results. The *DaS* algorithm generated linear outcomes, which were inaccurate but uniform. This consistency makes it possible to calculate the constant distance to the true value. The *EVOB* algorithm is comparable to *DaS* from its pattern. However, for the *EVOB* algorithm the difference to the real size value was always lower. The *Music* algorithm seemed acceptable for the first two distances, but turned out to be inaccurate for all other distances and lateral displacements.
Important: look at the source position first and if it is accurate, compare with results of max SPL area! If both are good, the algorithm is accurate.

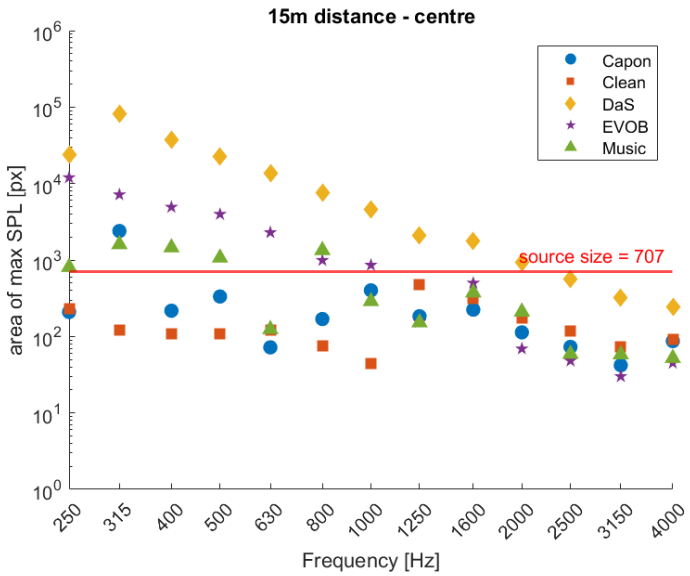


Figure 3.30: Calculated area of maximum SPL at a distance of 15m from the source. The source was placed in the centre.

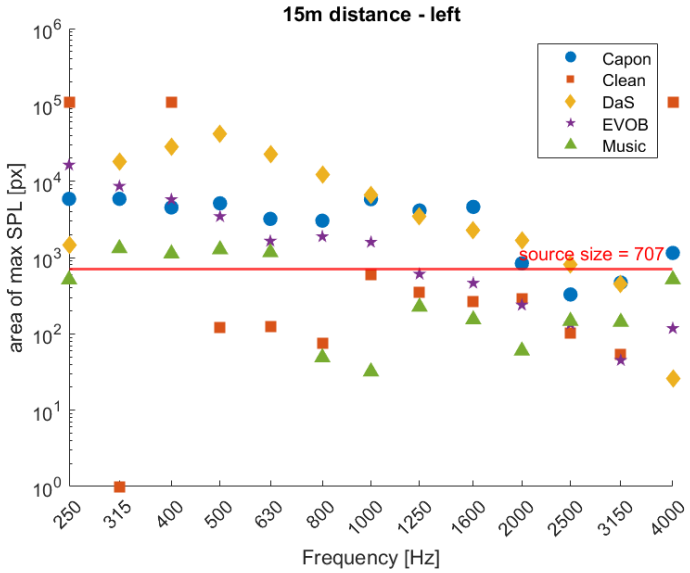


Figure 3.31: Calculated area of maximum SPL at a distance of 15m from the source. The source was placed to the left.

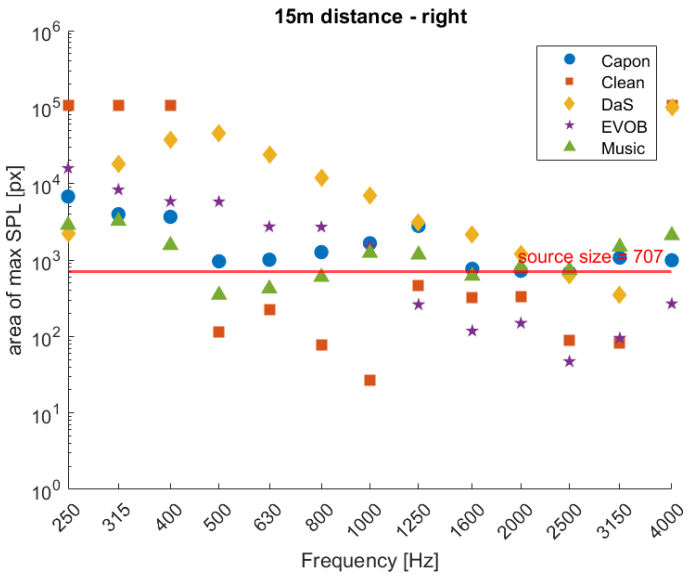


Figure 3.32: Calculated area of maximum SPL at a distance of 15m from the source. The source was placed to the right.

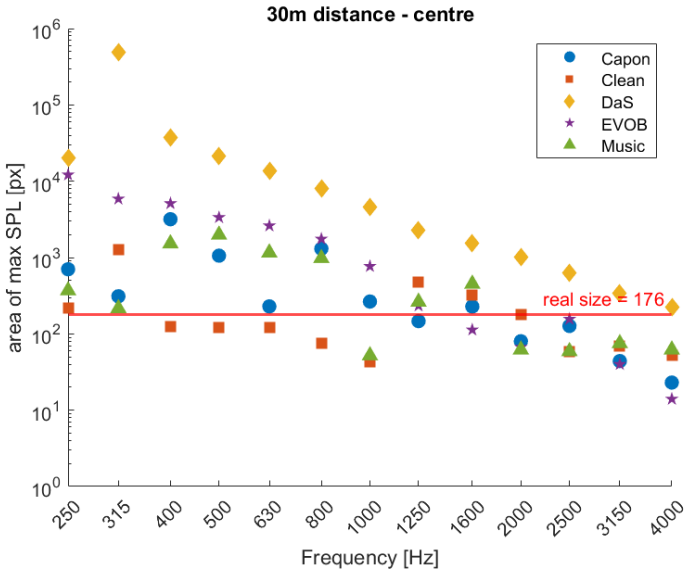


Figure 3.33: Calculated area of maximum SPL at a distance of 30m from the source. The source was placed in the centre.

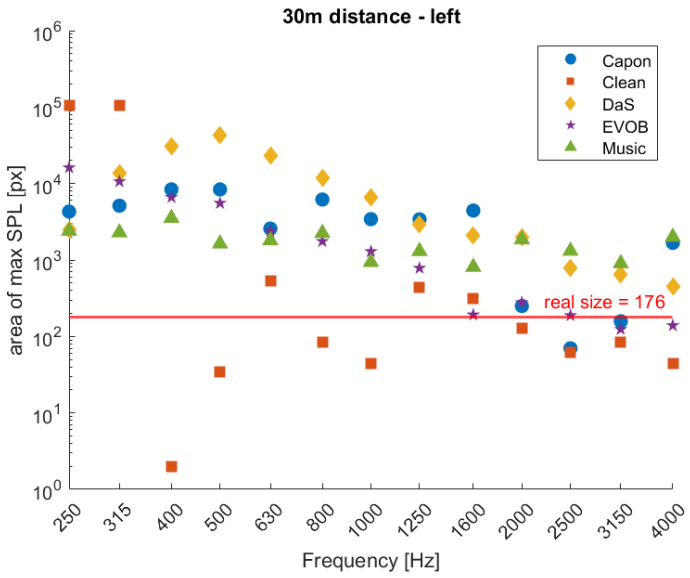


Figure 3.34: Calculated area of maximum SPL at a distance of 30m from the source. The source was placed to the left.

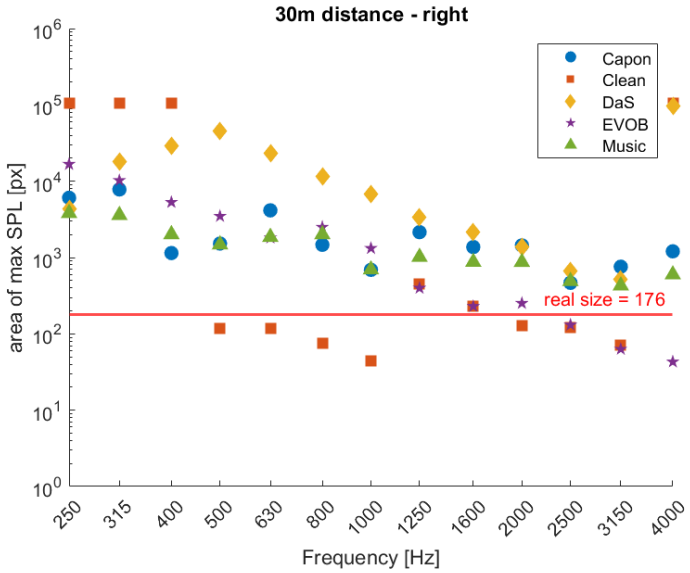


Figure 3.35: Calculated area of maximum SPL at a distance of 30m from the source. The source was placed to the right.

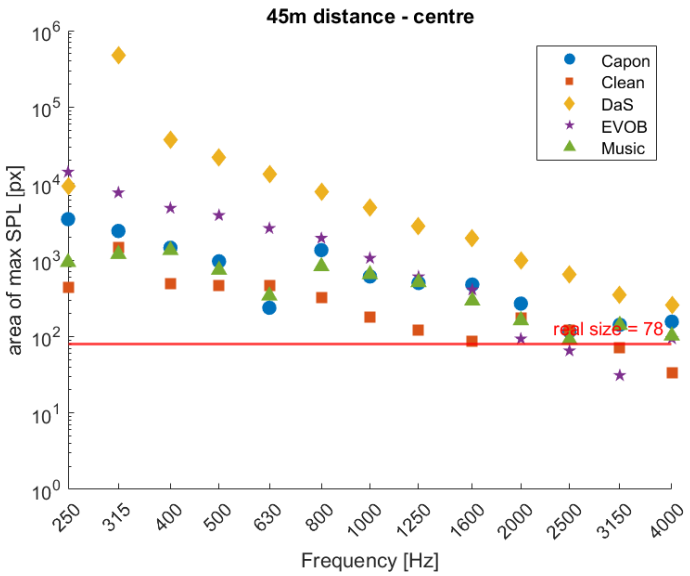


Figure 3.36: Calculated area of maximum SPL at a distance of 45m from the source. The source was placed at the centre.

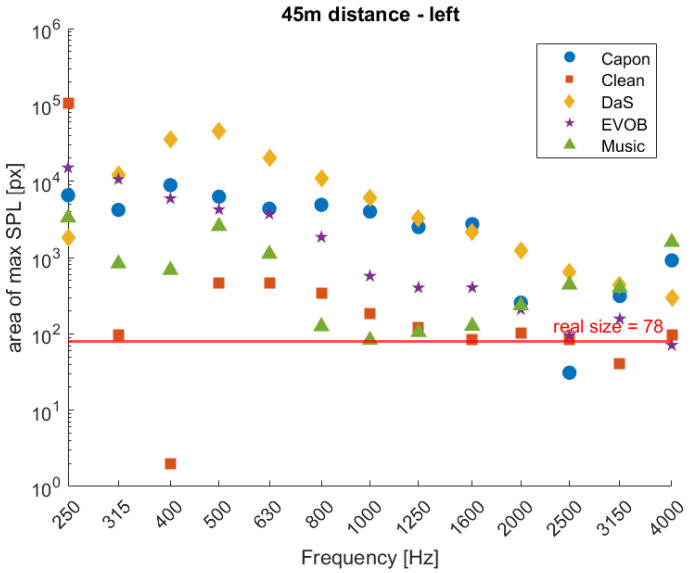


Figure 3.37: Calculated area of maximum SPL at a distance of 45m from the source. The source was placed to the left.

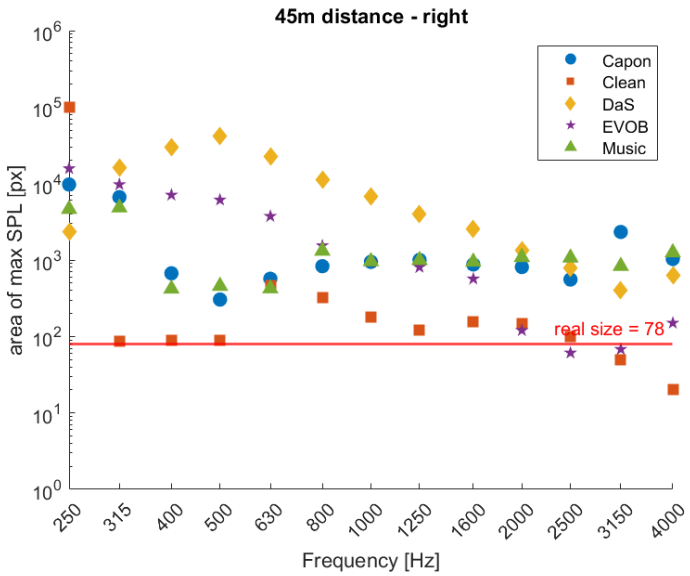


Figure 3.38: Calculated area of maximum SPL at a distance of 45m from the source. The source was placed to the right.

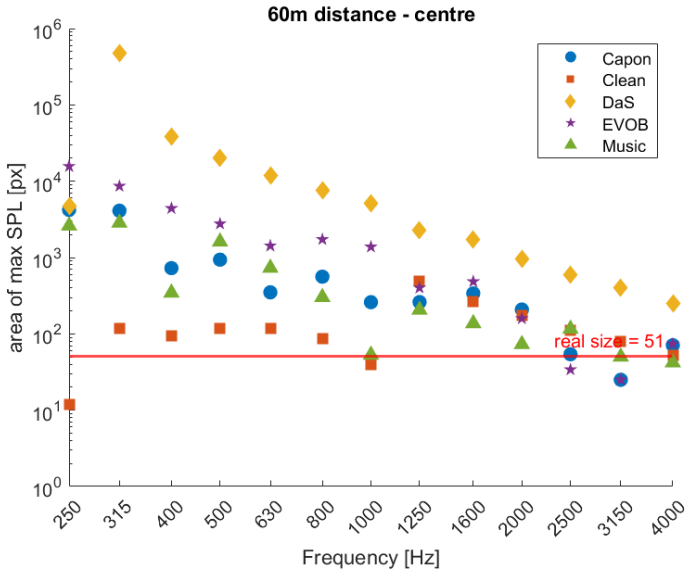


Figure 3.39: Calculated area of maximum SPL at a distance of 60m from the source. The source was placed at the centre.

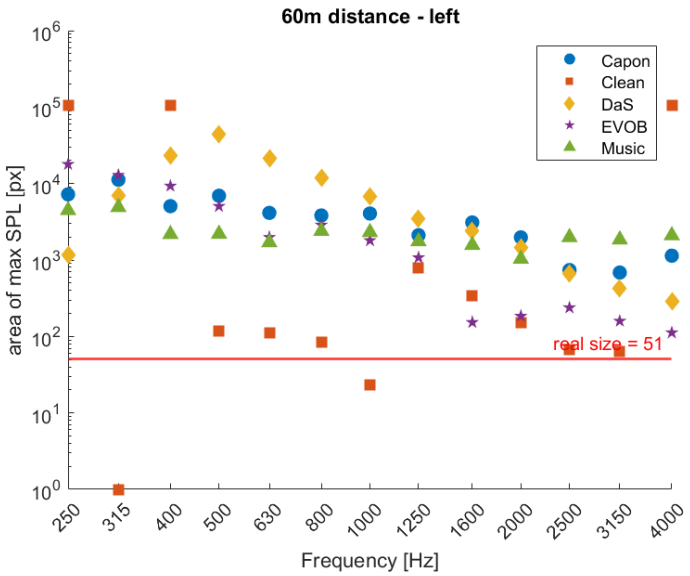


Figure 3.40: Calculated area of maximum SPL at a distance of 60m from the source. The source was placed to the left.

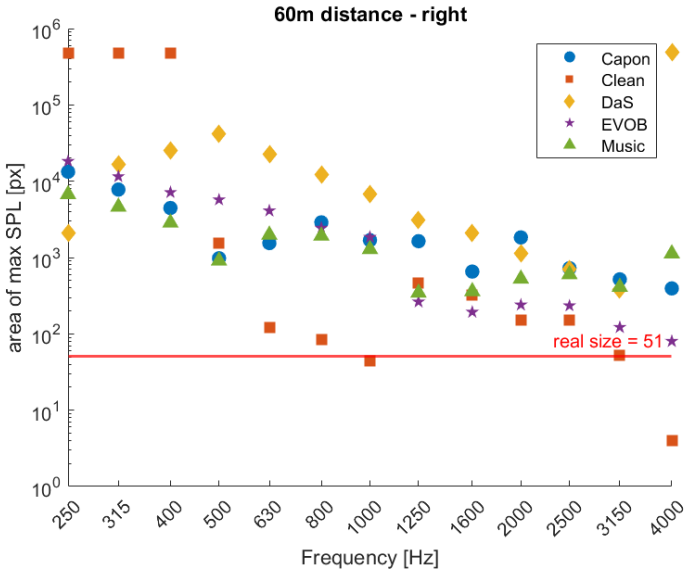


Figure 3.41: Calculated area of maximum SPL at a distance of 60m from the source. The source was placed to the right.

4 Conclusion and Future Work

In summary, the study provided overall insights into the strengths and weaknesses of five beamforming algorithms for specific sound source localization tasks. There are few but highly considerable aspects of this work that need to be acknowledged and which can lead to further research.

One of the most significant issues that emerged during this project was a fundamental error in the *CAE Noise Inspector* software and hardware combination. Specifically, it was observed that the *Music* algorithm used in the system tended to place the source of noise at the center of the image, regardless of other variables such as third octave frequency band, distance to the source, or orientation to the source. This error was reported by the research team to the development team of the *CAE Acoustic Camera*, and a new configuration file was subsequently sent to address the issue. However, the explanation provided was vague and did not specify the problem, and the closed-source nature of the software made it difficult to conduct a more thorough analysis of the algorithm's functionality. Therefore, all results obtained with the *Music* algorithm during this research were deemed insignificant. Unfortunately, the timing of the information sharing precluded the possibility of conducting a subsequent analysis using the corrected configuration file. Despite this setback, the research project identified an important issue that needs to be addressed in order to improve the accuracy and reliability of noise measurement systems.

The second element to consider is the inconsistency of the results. Although the conducted experiment adhered to strict and regulated modifications the obtained results are very irregular. Every change of the acoustic conditions was found to generate markedly divergent outcomes in the context of the algorithmic analysis being conducted. The acoustic conditions being defined as the power of the source, the input signal to the source and the sound field conditions around the source. Although the camera was tested in a controlled environment, no significant rules of conduct could be established. Especially after observing the algorithm's weak performance at low frequencies one can conclude that the complex sound field in an industrial harbour with its multiple, especially low-frequency sources such as cranes, lifting platforms, ship engines, turbines, trucks, pumps, trains and smaller machines presented the acoustic camera with a huge task. However, before drawing exclusively negative conclusions, it would be appropriate to investigate the behavior of the acoustic camera and its multiple functions in other variable environments. A possibility would be to

methodically analyse its functionality when multiple sources are active in its range of vision. By carefully analyzing the behavior of the system under these conditions, it may be possible to identify new patterns or trends that can be used to further optimize its performance or improve its design. Even so, conducting a methodical analysis of a system's functionality when multiple sources are active can be a complex and time-consuming process, requiring significant expertise in the relevant fields. Another area of interest could be to apply the various analysis steps that have been developed within this work to existing images of the measurements in the industrial harbor. This analysis would provide an opportunity to validate the accuracy and effectiveness of the algorithm in a real-world application. The industrial harbor as a complex environment, with unique challenges that must be overcome for the algorithm to be successful. Therefore, it would be beneficial to analyze the algorithm's performance in this setting and compare the results to those obtained from controlled experiments. By conducting this analysis, it may be possible to identify additional areas where the algorithm can be improved, optimized, or adapted to better suit the needs of the industrial harbor. Overall, reapplying the algorithm to the existing images of the measurement in the industrial harbor offers a valuable opportunity to assess the algorithm's performance in a practical setting and refine its capabilities for real-world applications.

Despite the growing body of evidence demonstrating the negative effects of noise pollution, there is still a lack of awareness regarding its consequences, such as stress, lack of sleep, and reduced concentration. To address this issue, the MON ACUMEN (Joint Urban-Maritime Active Noise Monitoring) project has been established with the aim of reducing noise impact in the commercial ports of the cooperation area, which includes Livorno, La Spezia, Cagliari and Bastia. To achieve this goal, the project aims to develop a common methodology for acoustic description analysis and noise detection, establish a shared design for monitoring systems and create a unified collection and verification process for collected data. These steps are crucial for effective planning, as required by European Directive *2002/49/EC*. By reducing noise pollution in these commercial ports, the project hopes to improve the quality of life and overall well-being of individuals living and working in the area. As part of the MON ACUMEN inter-regional project, noise maps with long term sound pressure level measurements, have been created, which can be combined with accompanying pictures of the origin of the noise. This approach offers a visually comprehensible representation of noise pollution that can sensitize even untrained readers to the issue. By visually demonstrating the sources and intensity of noise pollution, individuals may be more motivated to take action to reduce noise levels and improve their overall health and well-being. Overall, this could offer a promising strategy for reducing noise pollution in commercial ports and raising awareness of its negative effects on individuals and communities.