Localization of unmanned aerial vehicles with microphone arrays

Bence Csóka[1], Péter Fiala[2], Péter Rucz[3]

[1] Budapest University of Technology and Economics, E-Mail: csokabence996@gmail.com
[2] Budapest University of Technology and Economics, E-Mail: fiala@hit.bme.hu
[3] Budapest University of Technology and Economics, E-Mail: rucz@hit.bme.hu

Introduction

The topic of this paper is the automatized determination of the location and trajectory of sound sources by acoustic measurements with microphone arrays. It is possible to focus on a specific direction with the microphone array by amplifying and delaying the received signals in a suitable manner, with the help of beamforming algorithms. MUSIC (Multiple Signal Classification) is a relatively simple and fast beamforming algorithm that can be used to locate incoherent sound sources. This method is based on the separation of the received signal's cross-spectral matrix to signal and noise subspaces. In addition to determining the direction of arrival of the sound waves, we also estimated the exact position of the sound sources by having focal points at various distances from the microphone array, which is a novelty in the field of acoustics. We evaluated the performance of the MUSIC algorithm through simulations in the MATLAB environment. The method is suitable for the estimation of the direction of arrival of simulated sound waves, so its extension to track the position of the sources can be tested. Moreover, we used MUSIC for real measurements, where unmanned aerial vehicles act as sound sources.

MUSIC algorithm

There are two main problems that need to be solved to determine the position of sound sources: acoustical focusing and source localization. During focusing, we emphasize sound coming from a specific direction with the Delay and Sum method. The received signals of the microphones can be delayed and amplified separately in a suitable manner, so that the phase relations between the signals result in higher amplitudes after a superposition step (when focusing on the correct direction) [1]. The peaks of the delayed signals match up in time when the array focuses on the direction of the source, and thus the superposed signal has an even higher peak. Conversely, when the array focuses on a different direction, the peaks don't match up and the superposition makes weaken or completely negate each other, depending on the distance between the source and the focal point. The amplification and delay values for each microphone are determined by beamforming algorithms.

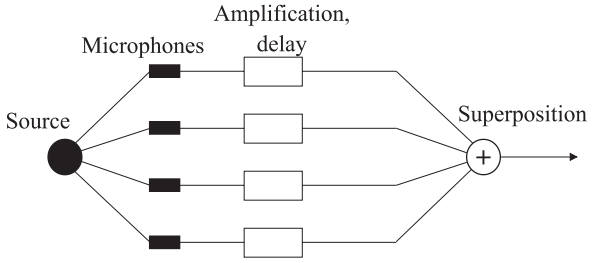


Figure 1: The Delay and Sum method.

The other problem, source localization is where beamforming algorithms become especially useful: here we estimate the position/direction of the source. A set of virtual source positions is selected in the space, which make up the acoustical canvas. From these points, we choose the one on which a placed virtual source generates a sound field that resembles the real one the most, and we assume that's where the real source is located.

These two tasks can be performed together as well: we can focus on every point of the canvas one-by-one, calculate an amplitude-like value for each of them with beamforming algorithms, then create an amplitude map, and the point with the highest amplitude is the estimated position of the source.

The microphones of the array can be placed in different formations. If the directions can be defined with only one angle, line formation is used. If two angles are necessary to define the directions, we can use cross, circle or spiral formations, instead of grid, which is likely too costly. The placement of the microphones influences the amplification and delay values applied on their received signals.

The beamforming algorithms are used in the frequency domain. This means that only the energy of a narrow band from the frequency spectrum of the received signals is used in the calculations: this energy value of the superposed received signals is the amplitude that is to be stored for every virtual source position. It is important to choose the middle frequency of this band correctly, because too low frequencies result in a blurred amplitude map, while at too high frequencies „phantom sources" can be observed, due to the violation of the principle of spatial sampling. The upper frequency limit for the spatial overlap can be calculated with equation (1):

$$f = \frac{c}{2d}. \tag{1}$$

MUSIC (short for Multiple Signal Classification) is a relatively simple and fast linear beamforming algorithm [2]. It separates the cross-spectral matrix of the received signals

into signal and noise subspaces, then uses the eigenvector of the noise subspace (notated with $\boldsymbol{U_n}$) and the sensing matrix (notated with **A**) for the calculation (equation (2)) [3][4][5][6]. The sensing matrix provides the link between the source distribution and the received signal vectors. The strength of MUSIC lies in its low computational cost and relatively high accuracy, but it can only be used with uncorrelated sources, and the number of sound sources must be estimated in advance.

$$P_{MUSIC} = \frac{1}{\boldsymbol{A}^H \boldsymbol{U_n} \boldsymbol{U_n}^H \boldsymbol{A}}. \quad (2)$$

Distance estimation

The virtual source positions of the acoustical canvas are usually placed along an imaginary flat or spherical surface. A canvas like this is suitable for direction estimation only. If we also want to estimate the distance (and thus the exact position) of sound sources, we must extend the canvas into three dimensions.

Changing the focal distance has a significant effect on the quality of the amplitude map. In figure 2, the results of a series of simulations can be seen, where I placed a sound source at a distance of 5 meters, and canvases with different distances from the plane of the microphone array, but they cover the same angle range. The closer the focal distance is to the actual source distance, the better the amplitude map.

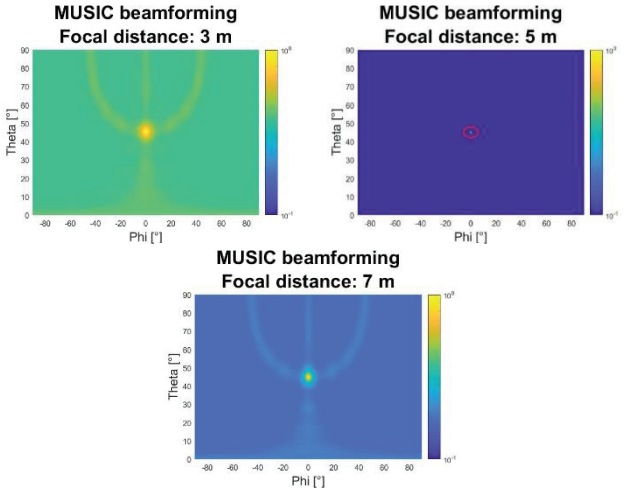


Figure 2: Beamforming with different focal distances.

Using different focal distances as a base idea, it is possible to devise a method for extending the acoustical canvas into three dimensions with distance estimation in mind. One such method is to make a direction estimation with a primary canvas the usual way, and along the estimated direction, we create a secondary canvas along a straight line, thus having virtual source positions with the same direction, but different distances. We apply beamforming with this new secondary canvas, and again, the point with the highest amplitude is the estimated position of the source.

Kalman-filter

The Kalman-filter is an algorithm that can estimate the state of systems that change with time. In this case, the system in question is a moving sound source, and the state is its position and velocity. The position coordinates can be extracted from beamforming algorithms. It starts from the standard state equation (equations (3) and (4)) of a discrete, linear, and time-invariant system [7][8]; together with the current measurements, it also takes the past states of the system into account, makes an a-priori estimation and then a correction step with the difference between the measured and estimated values.

$$\boldsymbol{x}(n+1) = \boldsymbol{A}\boldsymbol{x}(n) + \boldsymbol{B}\boldsymbol{u}(n) + \boldsymbol{w}(n), \quad (3)$$

$$\boldsymbol{y}(n) = \boldsymbol{C}\boldsymbol{x}(n) + \boldsymbol{D}\boldsymbol{u}(n) + \boldsymbol{v}(n). \quad (4)$$

The traditional Kalman-filter only gives an optimal estimation for linear systems, but it can be extended to work for nonlinear systems as well. One such extension is called „Unscented Kalman-filter" [8]. This version of the algorithm creates so called sigma-points around the state vector and makes the necessary calculation steps on them separately, thus countering the nonlinearity of the system. It can also be extended to be able to track multiple sound sources by increasing the size of the relevant matrices.

Simulations

The methods detailed above can be tested with simulations in the MATLAB environment. In two dimensions (where the directions can be defined with one angle or coordinate), the microphones are in a line formation, and the acoustical canvas lies along half of a circle in the same plane as the microphone array. Four point-sources are placed in the space that emit Gaussian white noises with different spreads. In these simulations, only the direction estimation is made.

In figure 3, the result of MUSIC beamforming is shown, on the left without background noise, on the right with a background noise applied, the signal-to-noise ration being 20 dB. The blue lines show the actual source directions, and the red line shows the estimation. In both cases, MUSIC gives a satisfactory result, with the peaks fitting on the blue lines fairly accurately, even with background noise applied (although the base estimation level is higher in this case, thus the peaks are lower compared to the base level).

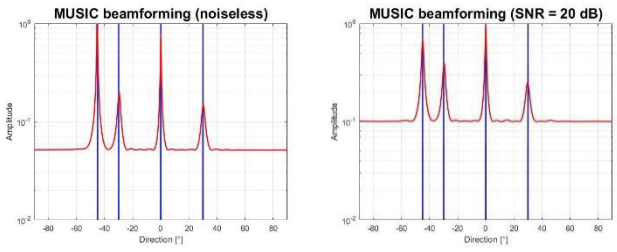


Figure 3: MUSIC beamforming in two dimensions, without (left) and with (right) background noise applied.

In three dimensions (where the directions can be defined with two angles) the microphones are in a cross formation, and the canvas lies along half of a sphere; the angles shall be interpreted as in a standard spherical coordinate system, with phi being the azimuthal, and theta the polar angle. Again, four point-sources are placed in the space that emit Gaussian

white noises with different spreads. The results of MUSIC beamforming in 3D are similar to the results in 2D (figure 4), even though the added dimension makes the direction estimation a bit more challenging, the algorithm creates a satisfactory amplitude map.

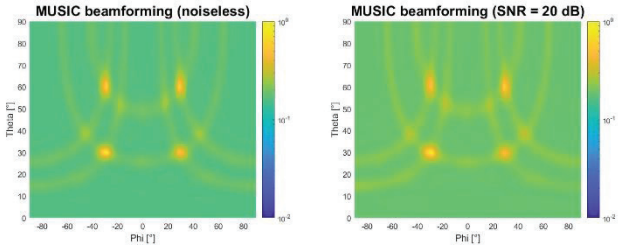
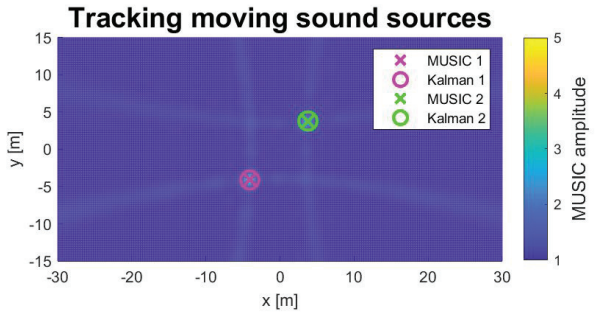


Figure 4: MUSIC beamforming in three dimensions, without (left) and with (right) background noise applied.

The final simulation presented in this paper demonstrates the algorithm extended with distance estimation (figure 4). Here the microphone array is in a cross formation, and the canvas lies along a rectangular are in a plane that is parallel to the plane of the array. There are two white noise emitting point sources in the space, one at 5 meters, the other at 50 meters from the plane of the array. The two sound sources are being tracked with both MUSIC (notated with crosses) and Kalman-filter (notated with circles). The sources are both moving along a straight line, and at one point, they are at the same direction from the centre of the array. Around this moment of the simulation, the algorithm cannot distinguish the two sources, and estimates an incorrect second direction, which makes the distance estimation incorrect as well. From the beginning and the end of the simulation (where there is a big enough difference in directions), we can conclude that it is easier to estimate the distance of sources that are closer to the array, because there are greater differences between the angles of incidence. We can also conclude that in this simulation, the distance estimation is mostly reliable (but with farther sources, not entirely accurate), if, and only if the direction estimation was correct as well. The Kalman-filter can reliably follow the MUSIC algorithm during the non-problematic section of the simulation. Depending on how we define the input parameters of the function that implements the filter, it can follow the beamforming algorithm more firmly or more loosely.



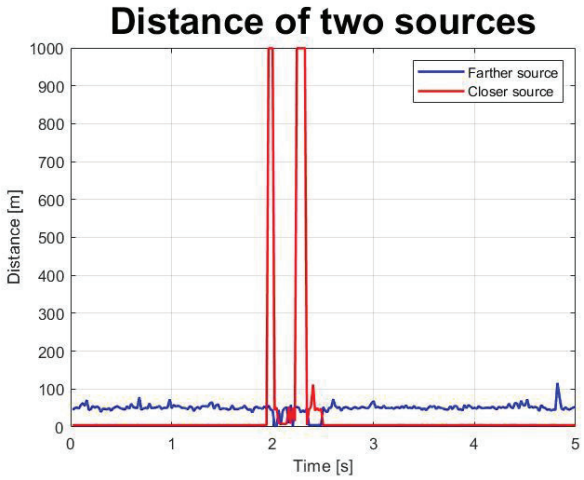


Figure 5: Direction and distance estimation of two moving sound sources.

Measurements

During our work, we were given the opportunity to take part in a measurement, where different groups were trying to determine the position of moving unmanned aerial vehicles (in other words drones) with different methods. One such method was the acoustical method.

The microphone array consisted of 48 microphones that were placed in a cross formation, stuck in holes on a wooden plate so that they remained firmly in position without movement. Several drones were moving along a similar trajectory, and below two short periods from the measurements are presented, with two different drones (one of them called Secopx8 and the other Tarot680).

The direction estimation with MUSIC and Kalman-filter gives overall somewhat worse, but still good results. The environmental conditions are less favourable here, disturbances are much less predictable. If the drones are loud enough and close enough to the microphones, the emitted sound stands out from the background. In this case, the direction estimation works reliably most of the time, except for a few moments when, presumably, strong background noise or ground reflection is received by the microphones. However, as the drones move farther from the array, the estimation gets less and less reliable.

In figure 6, the results of distance estimation for these two drones can be seen. The first picture shows Secopx8 during the timeframe when it moves from right to left in front of the array, and the second one shows Tarot680 when it is moving from far away and closer to the microphone array. In both cases, unlike during simulations, the gotten values are too unstable, change too much with time, thus the algorithm isn't currently capable of distance estimation with measurements, it must be developed further.

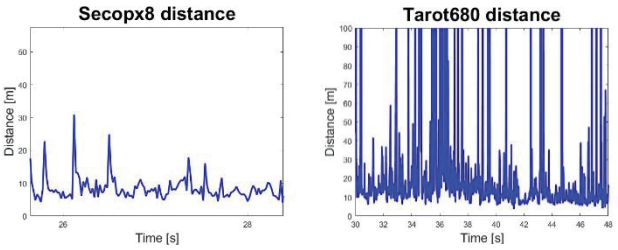


Figure 6: Distance estimation of two unmanned aerial vehicles.

Conclusion and future plans

In this paper, we discussed the determination of the position of sound sources with microphone arrays and beamforming algorithms. First, we gave a brief overview of the basics of beamforming and the Delay and Sum method. Then, we introduced a simple and fast beamforming algorithm called MUSIC. After that, we discussed a possible extension for the method to estimate not just the direction, but the distance of sound sources. We also touched on the subject of tracking moving sound sources with the Kalman-filter algorithm. Then, we moved on to presenting the results of simulations and measurements. We can conclude that MUSIC and Kalman-filter works reliably and accurately in simulations and in favourable measurement conditions, if only the direction estimation is concerned. However, the distance estimation only worked in simulations, and even there it was inaccurate for sources that were farther from the array.

Our work can be developed further in several different ways. We could test the algorithms with different sound sources that emit different signals, not just white noise. We could also simulate different microphone array formations, and assess their advantages and disadvantages compared to the cross formation. An important goal is to develop distance estimation further, increasing its accuracy and reducing its computational cost. If the algorithm can reliably and accurately determine the position of the sound source (or sources) and not just the direction, that opens the possibility of the estimation of velocity and sound power as well.

Literature

[1] J. Novoa, R. Mahu, A. Díaz, J. Wuth, R. Stern, N. B. Yoma: „Weighted delay-and-sum beamforming guided by visual tracking for humanrobot interaction”. 2019, arXiv:1906.07298.

[2] R. Schmidt: „Multiple emitter location and signal parameter estimation”. IEEE Transactions on Antennas and Propagation Vol. 34, 1986, pp. 276–280.

[3] A. Xenaki, P. Gerstoft, K. Mosegaard: „Compressive beamforming”. The Journal of the Acoustical Society of America, Vol. 136 (1), 2014, pp. 260-271.

[4] M. Mohanna, M. L. Rabeh, E. M. Zieur, S. Hekala: „Optimization of MUSIC algorithm for angle of arrival estimation in wireless communications”. NRIAG Journal of Astronomy and Geophysics, Vol. 2 (1), June 2013, pp. 116-124.

[5] Q. Zhao, W. Liang: „A Modified MUSIC Algorithm Based on Eigen Space”. In: Jin D., Lin S. (eds) Advances in Computer Science, Intelligent System and Environment. Advances in Intelligent and Soft Computing, Vol 104. Springer, Berlin, Heidelberg, 2011, doi: 10.1007/978-3-642-23777-5_45.

[6] P. Gupta, S. P. Kar: „MUSIC and improved MUSIC algorithm to estimate direction of arrival”. 2015 International Conference on Communications and Signal Processing (ICCSP), Melmaruvathur, 2015, pp. 0757-0761, doi: 10.1109/ICCSP.2015.7322593.

[7] D. Simon: „Optimal State Estimation – Kalman, H∞, and Nonlinear Approaches”, John Wiley & Sons, Inc., Hoboken, New Jersey (2006).

[8] Belső Zoltán, Gáti Balázs, Koller István, Rucz Péter, Turóczi Antal: "Design of a nonlinear state estimator for navigation of autonomous aerial vehicles" Repüléstudományi közlemények (Aviation scientific publications) XXVII/3 pp. 255–276 (2015).