

Chapter 4

Evaluation

In this chapter, experimental results are presented while setting focus on different aspects. At first, the results of the methods on single robots must be investigated before evaluating the performance of the final localization algorithm. Thus, the results of the direction detection methods is examined profoundly on one robot taking one measurement as example in section 4.1. There, the focus is set on the TDOA methods itself for each channel pair before combining them to a robot direction result.

In chapter 3, the presence of determinable supplementary information was discussed regarding the SNR, PSNR and distance approximation. If and in what extent these factors are applicable for the source localization is analysed in section 4.2.

To investigate the whistle sound source localization of the robots as team, measurements were done with five Naos on the field of the HULKs' laboratory as specified in section 4.3.1. The results of the TDOA methods for all measurements are presented and compared in section 4.3 to have a closer look at their performance. Depending on the accuracy of the individual direction results, the quality of the team filter is limited. Therefore, the angular error of the single robots is demonstrated for each measurement and method in addition.

Primary, whistle sounds were recorded in the laboratory of the HULKs. If not emphasized particularly, the height of the sound source is 1.5m.

4.1 Single Robot Method Comparison

For comparability of the results, one exemplary measurement is utilized to present and analyse the TDOA methods. In this data set, the sound source is placed at the right front of the robot with 4.5m distance. This corresponds to an angle of -33.7° in robot coordinates. Some samples around the signal start of the received signal data is plotted in 4.1 for all channels.

As the next sections focus on the performance of the TDOA methods, the start index is set manually. A frame size is defined with 256 samples and is selected around the start index for the correlation methods as explained in section 3.3. For the phase method, the first frame with a size of 64 samples is chosen where a whistle is detected in all channels.

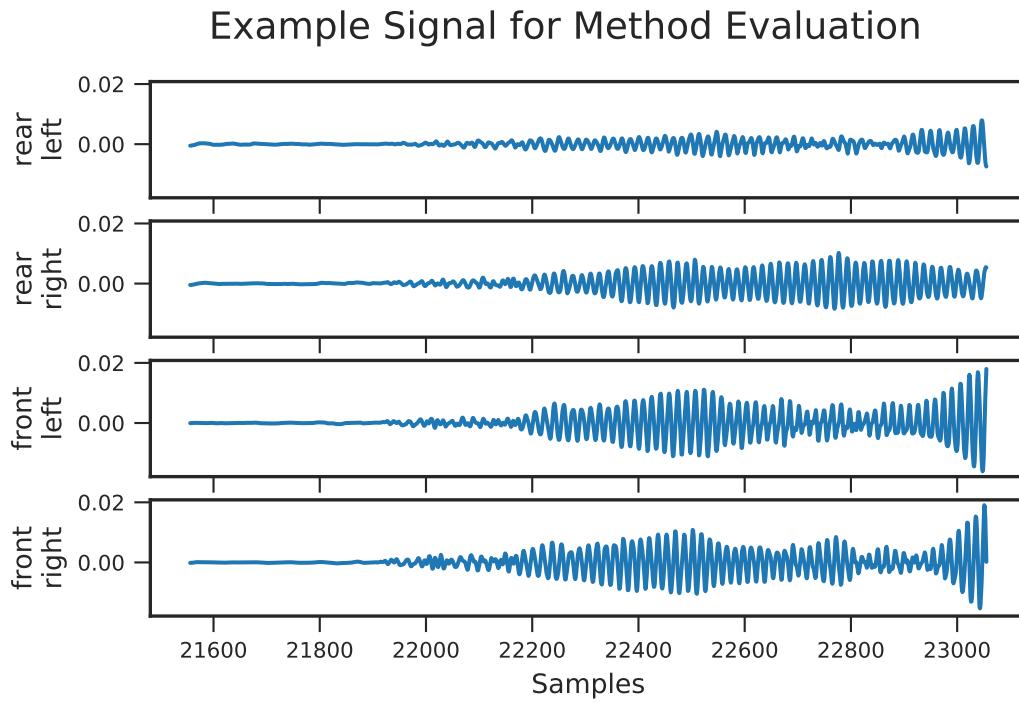


Figure 4.1: Signal start section of a whistle sound recorded from front right.

In the following, the correlation function $R_{x_a x_b}$ of two signals x_a and x_b as in chapter 2 is shortened to R_{ab} for simplicity.

4.1.1 Cross Correlation

To visualize the result of the CC, the correlations are plotted in fig. 4.2. For R_{23} and R_{13} the peak is clearly traceable. However, for the other CC the problem of a low maximum peak arises as mentioned in section 2.3.

Base Channel	Next Channel	Delay	Candidate (-)	Candidate (+)
0	1	-8.25	-144.9	-35.1
1	3	-4.59	-17.4	78.6
2	0	9.16	-30.6	-30.6
3	2	3.94	-150.2	-29.8

Table 4.1: Cross correlation delay results of singal from front right.

According to the delays in table 4.1, the final result of the CC is -26.9° which emerges an error of 6.8° . The delay between channel 2 and 0 is larger than the maximal delay of 6.85 samples and therefore cut to the maximal sample delay. Besides these, the TDOA between the channel pairs produce one appropriate direction candidate which correctly points to the sound source.

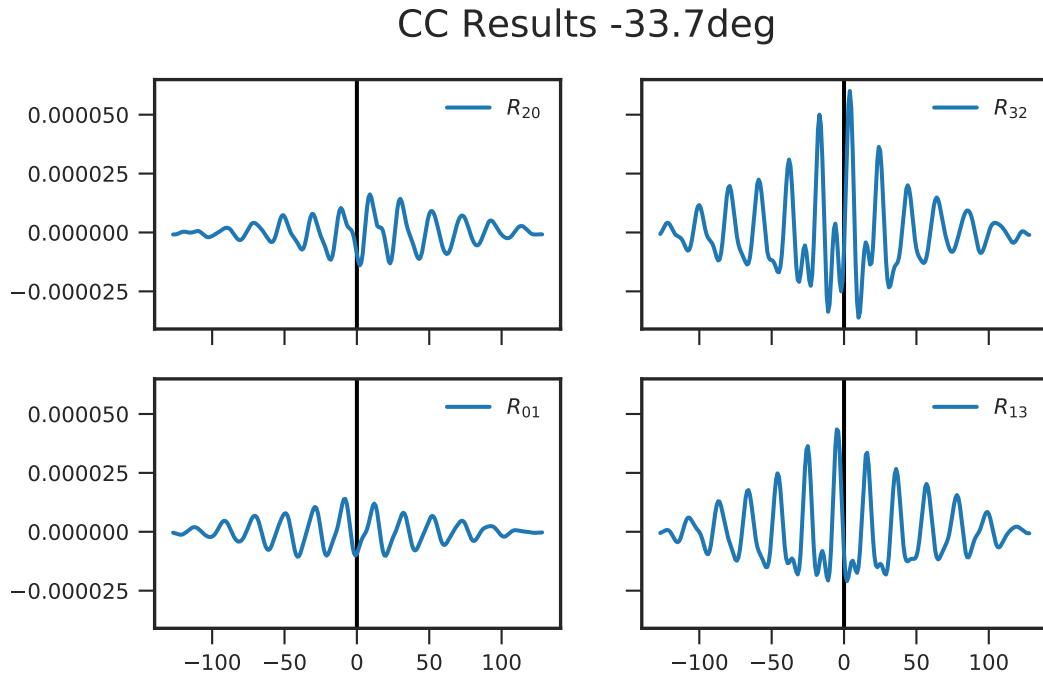


Figure 4.2: Cross correlation results of signal from front right (-33,7°).

4.1.2 Generalized Cross Correlation

Figure 4.3 presents the cross correlation result by the GCC method of the same signal data as in the previous section. The subsample delays for each channel pair and their resulting direction candidates are listed in table 4.2. From this, a final direction of -30.0° is determined resulting in an error of 3.69° . It is apparent that the peaks of the GCC are better to detect than the peaks of the CC.

Base Channel	Next Channel	Delay	Candidate (-)	Candidate (+)
0	1	-8,28	-144,7	-35,3
1	3	-4,09	-22,8	84,0
2	0	7,60	-30,6	-30,6
3	2	4,13	-148,7	-31,3

Table 4.2: Generalized cross correlation delay results of singal from front right.

4.1.3 Phase Difference

For detecting the source direction with phase difference, a smaller frame size of 64 samples is defined. Previously in section 3.3, two cases were introduced where either a fixed frequency is set or a frequency that is dominant for all four channels is taken for reference.

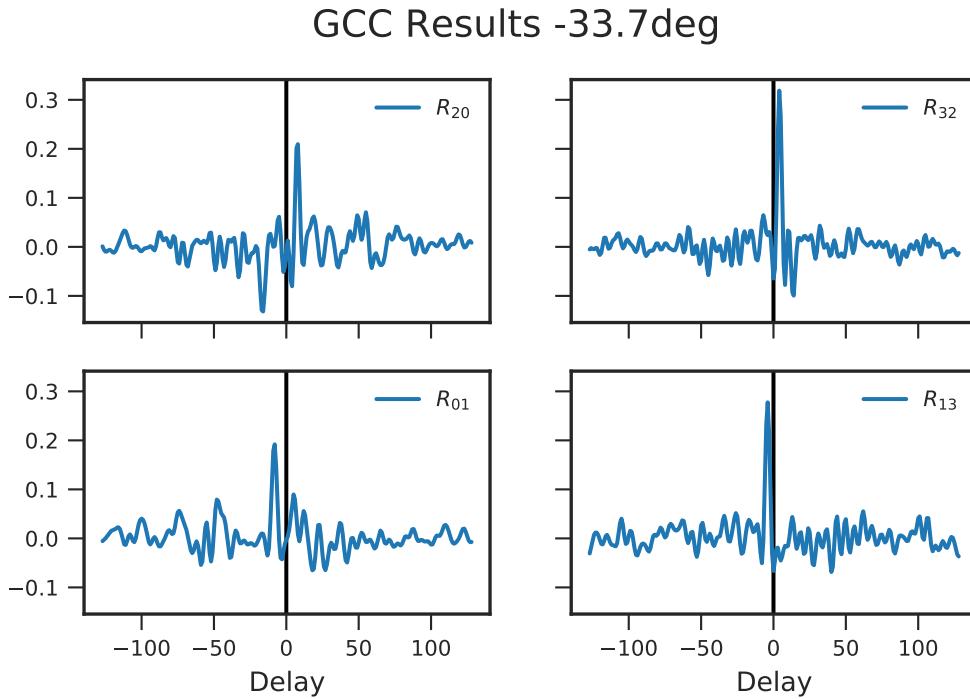


Figure 4.3: Generalized cross correlation results of signal from front right.

Dynamic Frequency Selection

First, the result of the dynamically selected frequency is presented. As stated in the implementation chapter, the frame is chosen where the frequencies of the maximal amplitudes coincides for all channel which is at 2756,25Hz. In the upper plot of fig. 4.4 one sees the received microphone data which will be Hann windowed and then transformed into frequency domain by FFT. The resulting phases and amplitudes are listed in table 4.3. For comprehensibility, the determined frequency information visualized by wave signals with the detected phases and amplitudes in the lower subplot of fig. 4.4. Due to the larger distance between channels 0 and 1, the phase difference information must be neglected because the phase difference is ambiguous. Outcome from the applied phase differences is $-29,2^\circ$ by the combination of $-17,6^\circ$, $-30,6^\circ$ and $-39,3^\circ$.

Channel	Phase [deg]	Amplitude
0	-1,55	0,00144
1	-177,7	0,00287
2	173,4	0,00279
3	-75,0	0,00372

Table 4.3: Phase and amplitude of frame signals with $f_c = 2756,25\text{Hz}$.

Base Channel	Next Channel	Phase Difference [deg]	Candidate (-) [deg]	Candidate (+) [deg]
1	3	-102,7	-17,6	78,8
2	0	173,4	-30,6	-30,6
3	2	113,1	-140,7	-39,3

Table 4.4: Phase differences and resulting direction candidates of example data with phase method.

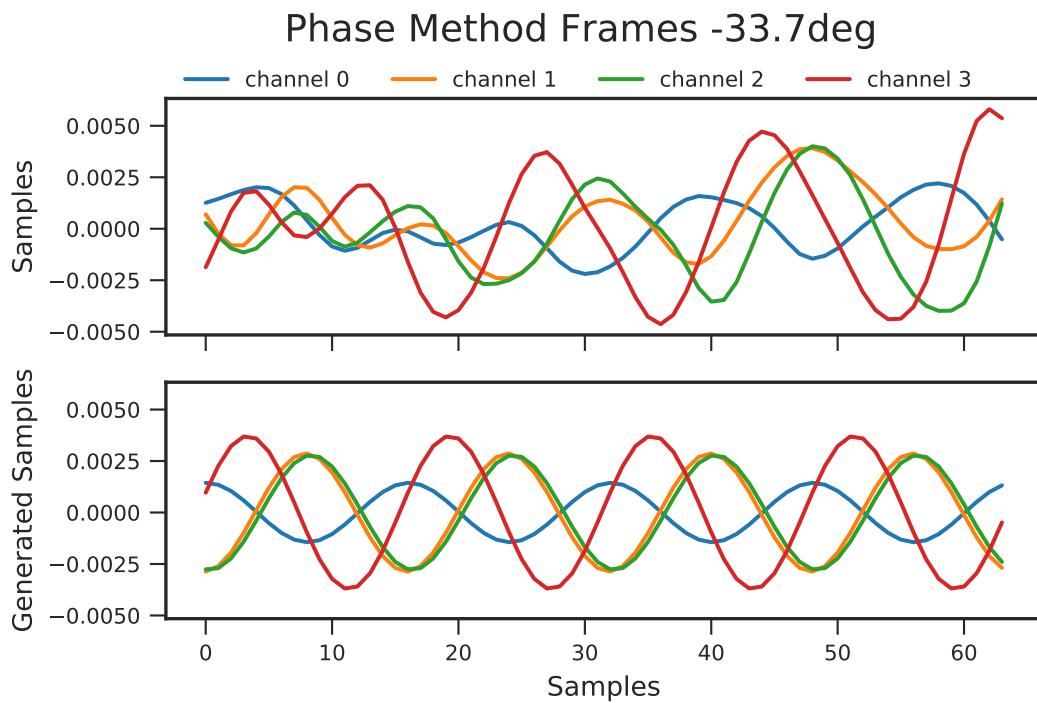


Figure 4.4: Frames used for the direction detection by phase method.

Static Frequency Selection

Secondly, the frequency to examine f_c is set to the first represented frequency larger than 2600Hz which is 2627,1Hz with a FFT length of 256. At this frequency, outcome of the direction candidates listed in table 4.5 is $-29,6^\circ$ which results in an error of $4,1^\circ$.

Base Channel	Next Channel	Phase Difference [deg]	Candidate (-) [deg]	Candidate (+) [deg]
1	3	-79,1	-26,8	88,0
2	0	167,7	-30,6	-30,6
3	2	88,5	-148,7	-31,3

Table 4.5: Resulting candidates of phase difference method with fixed frequency 2627,1Hz of example measurement from front right ($-33,7^\circ$).

Static Frequency Value

In order to observe the influence of the chosen frequency f_c the phase method is tested with different frequencies in the whistle range. To have a more generalized conclusion, more measurements are taken for this investigation. From eleven positions, all recordings of the whistle sounds are done with one robot arranged at the center point. This data corresponds to the measurements of the robot with number 26 in section 4.3.1.

As shown in fig. 4.5 shows, the Root Mean Squared Error (RMSE) is high for frequencies smaller than 2600Hz. With a frequency of 2024.12Hz, error is largest. The result complies with the information in fig. 3.4 showing that in a whistle signal frequencies higher than 2500Hz are dominant. With this outcome, the fixed frequency is set to 2627,1Hz for further usage of the direction detection by phase method. Limitation exists due to the ambiguity of the signal which is content of section 3.3.

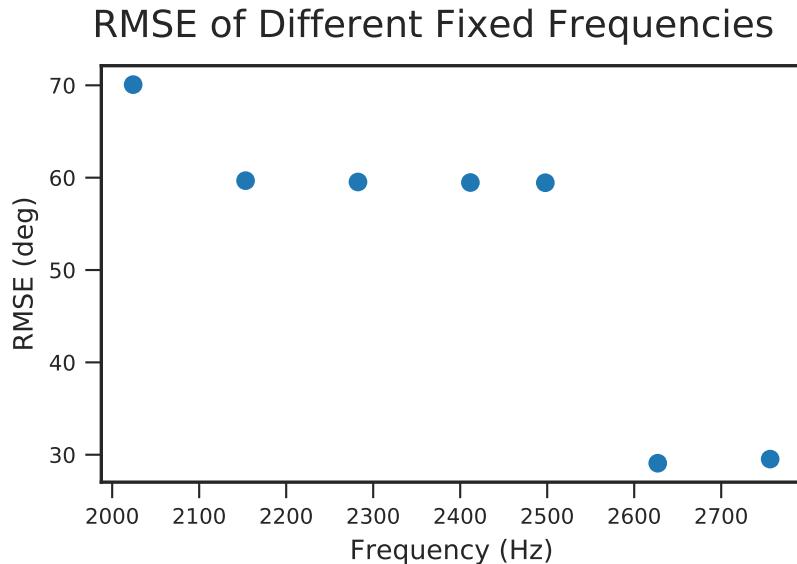


Figure 4.5: Result of all measurements done with robot 26 to compare different fixed frequency values in whistle range.

Frame Number

Not only does the frequency play a major role for the phase method, but also the chosen frame. To evaluate if and how the result changes over time, the frame to utilize is shifted by half the frame size for all measurements of section 4.3.1 for the robot at the center point. In fig. 4.6 one sees that the first channel frames with zero shift gives the best result with a RMSE of 29,1°. It validates that not any frame can be chosen for the direction detection what again confirms that the signal start detection plays a large role for the correct result.

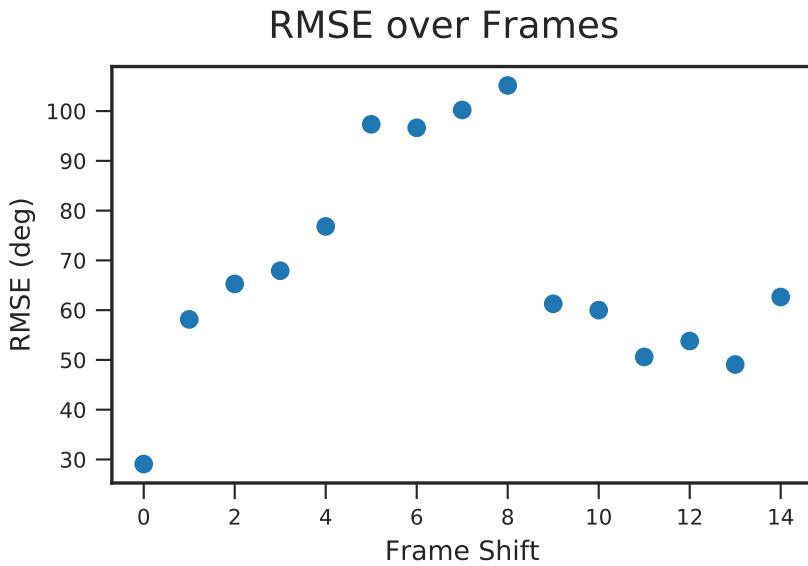


Figure 4.6: Changing results by shifting the frame over the samples. All measurements of section 4.3.1 with robot nr. 26 are taken into consideration.

4.2 Additional Information

On top of the direction detection by TDOA, additional information can be won by received microphone data. This information can be used to improve the single robot result or to feed the team filter with beta information. As section 3.4 has described, the distance of the sound source can be estimated approximately for nearby signals that are aligned with the x-axis of the robot's head. Another intuitive approach is the inspection of the SNR which is assumed as higher for closer sources. Apart from this, the PSNR of the GCC defined in section 3.6 is investigated.

4.2.1 Distance Approximation

Section 3.4 clarified the conditions to identify the distance of the sound source by one robot. To examine the validity of this statement, measurements from the front and back of the robot are collected and evaluated. Table 4.6 lists the distances and their results by the all methods. For both measurements with zero distance, the orientation of the whistle differed. 180° indicates that the whistle was turned in the opposite direction of the robot. In the other case (0°), the whistle was aligned with the robot. The distance is represented in robot coordinates, so that positive distance expresses that the source was placed in front of the robot and oriented towards it and vice versa.

As the results show, the distance can be approximated with adequate error. One can see that the GCC results are erroneous for small distances, but gives a correct approximation steadily. Compared to this, the CC method performs better for small distances, but fails completely for some measurements. Those failing cases mostly arose when lateral delays exceeded the maximum lateral samples incorrectly. The phase methods provides most incorrect results. Especially measurement 10 stands out by being double the real value.

For all measurements, the resulting direction angle of the sound source were accurate. Right-

Nr.	Real Distance [m]	GCC Result [m]	CC Result [m]	Phase Result [m]
1	0,9	∞	∞	∞
2	0,6	∞	∞	∞
3	0,3	0,35	0,25	0,13
4	0,0 (180°)	-0,13	-0,15	-0,23
5	-0,0 (0°)	0,22	0,21	0,02
6	-0,3	-0,15	-0,27	-0,45
7	-0,6	-0,34	-0,50	-0,62
8	-0,9	-0,70	-0,91	-0,99
9	-1,2	-1,00	-1,28	-1,71
10	-1,5	-1,39	-1,59	-2,98
11	-1,8	-1,72	-2,07	-3,33
12	-2,1	-2,16	∞	-3,02
13	-2,4	-2,31	∞	∞
14	-3,75	-3,66	-9,51	-4,15
15	-6,4	-7,35	-7,27	∞
16	-9,8	∞	∞	∞

Table 4.6: Result of front and rear distance for all methods..

fully, the algorithm detects sources that are out of observable range.

Unfortunately, for real cases one can not rely on the height parameter of the sound source which varies from the referee's body height. Having this as approximation only, the distance result should be handled with care.

4.2.2 SNR

The validity of a result can be brought in the Bayesian updating algorithm. Depending on the uncertainty, the covariance value of the incoming result can be adjusted so that unreliable results have less influence to the filtering. If a relation between received signal strength and distance to the source exists, it would be a simple way to predict the sound source position roughly. This can help for example, when team filter intersections are clustered so that multiple location results exist.

Taking the measurements of 4.3.1, this hypothesis is investigated by looking at the relation between distance and relative SNR. By reason of the whistle not blown equal for all measurements, the SNR is scaled by the overall mean of all robots' SNRs for one measurement. In fig. 4.7 no straightforward link between both values can be found unexpectedly.

Due to the result, further analysis on the SNR values on individual robots is done for evaluation of the microphones' validity. For this purpose, a signal is played back digitally from fixed distance and constant volume with different angles. The main purpose of this measurement is to ensure that the lone channels are not biased. 14 measurements of a 3kHz sine signal with

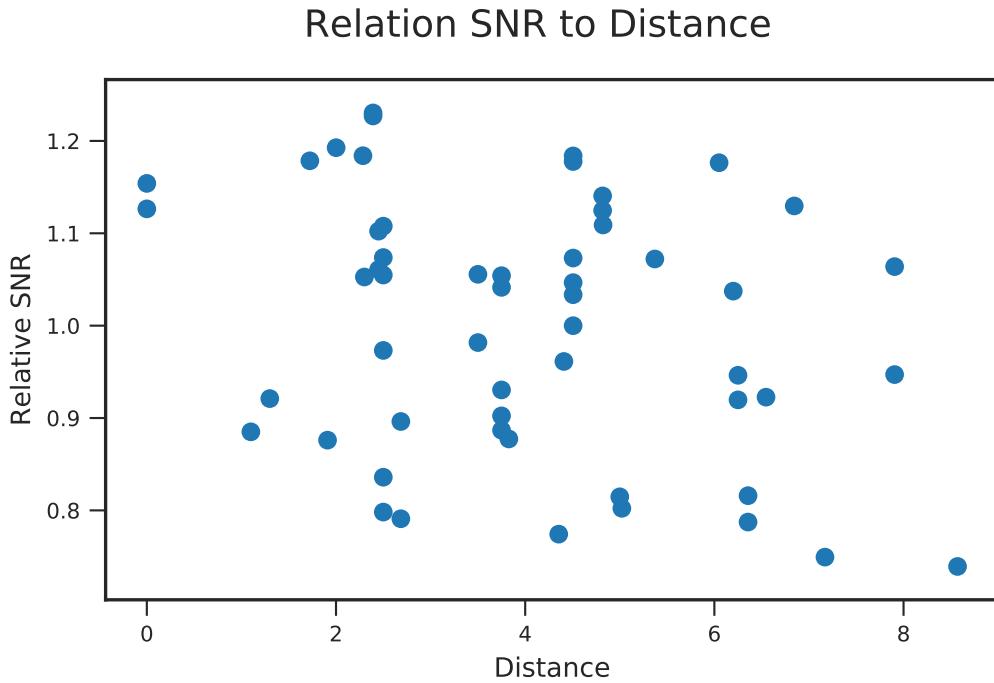


Figure 4.7: Visualization of relation between SNR and distance.

are distance of 0.73m are done but no tendency is detectable. Further evaluation is done by determining the channel with the maximum SNR of one measurement. It is expected that the nearest channel to the sound source has maximum SNR. At 85,71% of the cases this assertion could be evidenced. From this, it can be said that the general recordings of the microphones are neither biased nor falsified.

The same procedure is done with the real whistle recordings of the measurement in 4.3.1. Using these measurements, only 54,55% of the maximum SNRs match with the expected channels. Consequentially, one must assume that the environmental circumstances like multi-path propagation and reflection have large influence on the signal data. Thus, neither for a single robot direction result nor for the team filter the SNR can be utilized.

4.2.3 PSNR

As referred in section 2.4, the main characteristic of the GCC-PHAT algorithm is the emerging sharp peak. In conclusion, one can assume that the lack of a sharp peak indicates a less valid delay result of the GCC.

Informative Value

This statement is examined by comparing the PSNR value to the error of the direction angle resulting from the GCC-PHAT delay. In the following, the PSNR is recorded as high if it exceeds 17,5 whereas the PSNR value ranges from 10,1 to 28,8 for measurements in section 4.3.1.

Firstly, it is looked at each channel pair by determining the error between true direction and the most suited direction candidate emerging from the GCC delay. The results are split by the PSNR value. If the PSNR of the GCC is greater than the threshold of 17,5, the RMSE of its result is grouped to the errors with high PSNR. Elsewise it pertains to the errors with low PSNR. This valuation is done with the measurements of section 4.3.1 and manually set start indexes. Having 76 correlations assessed with low PSNR, the RMSE of this group is $35,77^\circ$. Compared to this, the RMSE of the remaining 144 measurements is $15,86^\circ$ only.

To see the impact on a complete robot result, the same is done with the final angular errors of single robots results. Here, the RMSE of the lower PSNR case results in $25,36^\circ$ whereas the error of the other case is $14,41^\circ$. From this, one can identify the PSNR as valid enriching information.

Frame Selection

Another perspective is to include the PSNR information into the single robot whistle direction determination. In fig. 4.8, the frame to investigate is shifted before and after the signal start index. Here, the same measurement as in section 4.1 is used where the robot was positioned at the center point while the whistle is blown at $-33,7^\circ$ with 4,5m distance. The frame size of the GCC is set to 256 samples and the shift was set to a quarter of the frame size. Samples of the rear left channel are plotted for better understanding in the upper graph of fig. 4.8. The second graph shows the RMSE of the robot direction result over the frame shifts. For the lower graph, the mean over the PSNRs of all channels is presented. One sees, how the error is low with high PSNR what confirms the previous statement in this section. For shifts smaller than -2, the signal is not represented in the frames yet what explains the high error. An important notice is that the result changes with proceeding frames what indicates that the implemented GCC-PHAT method is not suitable for arbitrary frames.

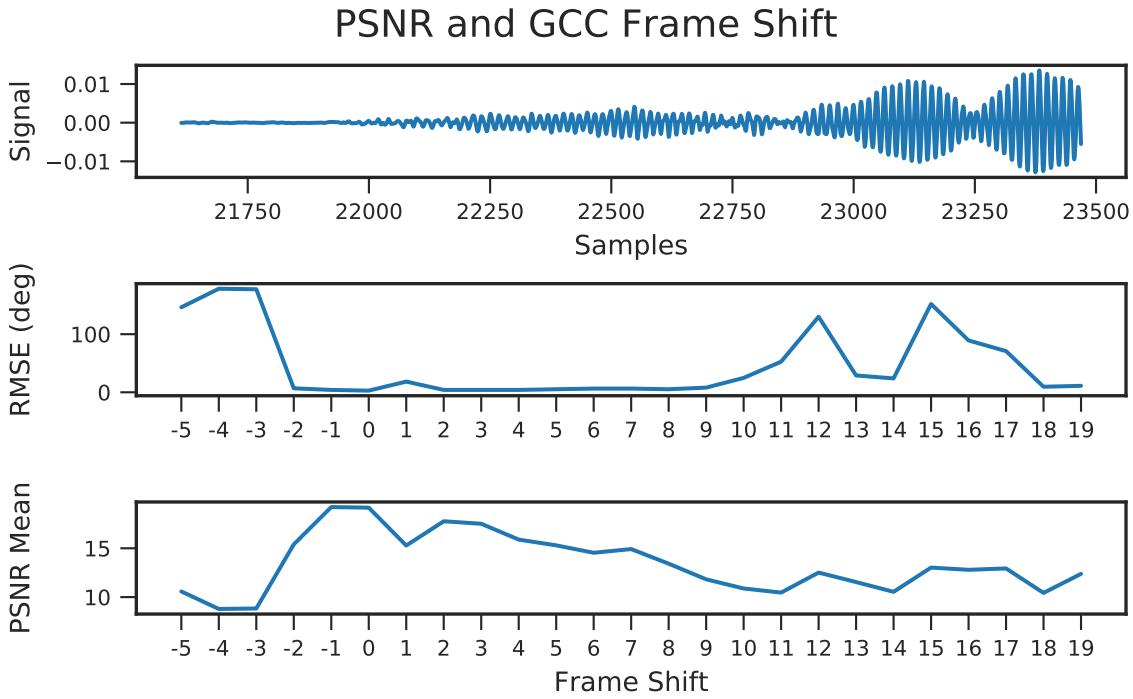


Figure 4.8: Relation between PSNR and selection of the frame in time. Signal data of the rear left channel is plotted in the upper window. In this measurement, the whistle is positioned at right front of the robot.

4.3 Team Evaluation

In the last chapters the performance of the three TDOA methods were validated on standalone robots independently. Here, the focus is on the final whistle localization algorithm as multi agent system. With the results of the single operations as input, the team whistle localization outputs an absolute sound source position. To provide a decoupled result to the signal start detection, the start indexes were set manually.

The size of the field used in this work is smaller than the regular Standard Platform League (SPL) field with 7.5m length and 5m width. Further information about the measurement setup is introduced in section 4.3.1.

4.3.1 Measurement Setup

11 measurement were taken with five robots on the small-sized SPL field of the HULKs laboratory. Figure 4.9 illustrates the positions of the Nao robots and the positions of the whistle sound sources. According to these, the x and y values of the sources and robots are listed in table 4.7 and table 4.8. The orientation θ of the robots are defined relatively to the x-axis and correspond to the definition in section 3.1.1. In the following sections, the signal data from these measurements will be used mainly.

Nao	x [m]	y [m]	θ [deg]
21	3,75	2,5	-40,2
24	3,75	-2,5	90
26	0	0	0
27	-3,75	-2,5	66,06
28	-2,45	0	0

Table 4.7: Positions of the robots for evaluation measurement.

Measurement	x [m]	y [m]
[0]	3,75	2,5
[1]	3,75	-2,5
[2]	-3,75	-2,5
[3,9]	-3,75	2,5
[4]	-2,45	0
[5]	2,45	0
[6,10]	0	0
[7]	0	-2,5
[8]	-6.05	0

Table 4.8: Positions of the whistle sound sources for evaluation measurement.

4.3.2 GCC Method

To determine an overall result, each robot executes the sound source direction detection with the GCC-PHAT method standing alone. After that, the results are input into the team decision filter as specified in section 3.7 which determines a final sound source position result with Bayesian updating.

For further clarification, details of the result for measurement 1 of section 4.3.1 is presented extensively here. Figure 4.10 illustrates the result of the relative direction angle outputs γ of the individual robots listed in table 4.9. Same as in fig. 4.9, the yellow dots symbolizes the robots with its orientation as short yellow line. The resulting arrows represent the direction γ where the sound source is predicted. In the graphic the real source is marked as star and the resulting position as cross.

The final result and its corresponding errors are listed table 4.11. With an absolute distance error of less than 1m and small angular error, the source localization works sufficiently for this case of application.

Table 4.11 shows the distance and angle errors for all measurements in section 4.3.1. The RMSE in distance being 0.87m and angular RMSE being $5,07^\circ$ one can say that the GCC-PHAT algorithm works well for whistle sound source localization.

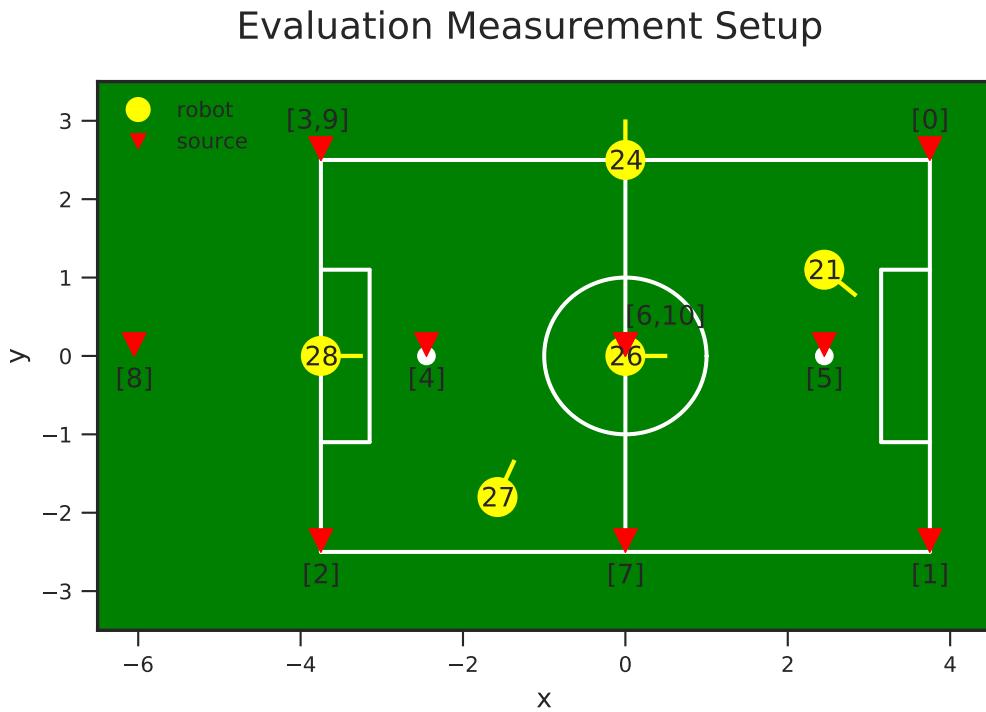


Figure 4.9: Setup of robots and sound source positions for the evaluation measurement.

Nao	γ [deg]	Abs. Error [deg]	Mean PSNR
21	-26,22	3,71	18,3
24	-133,77	9,32	16,8
26	-30,19	3,50	19,6
27	-75,26	1,71	17,4
28	-15,90	2,53	15,1

Table 4.9: Resulting directions of the single robots with GCC-PHAT method for a whistle sound signal in the right front corner of the playing field.

	Result	Error
Position x [m]	3,38	-0,37
Position y [m]	-1,85	0,65
Angle	33,18°	1,57°
Distance [m]	3,85	0,74

Table 4.10: Whistle localization result of measurement 1 with GCC-PHAT method.

4.3.3 CC Method

Evaluation is done of the same measurements in section 4.3.1 with normal CC. The results show that the normal CC without weighting performs poorer compared to the GCC-PHAT method with an RMSE of 1,45m in distance and 10,67° angular. Indeed, the statements in

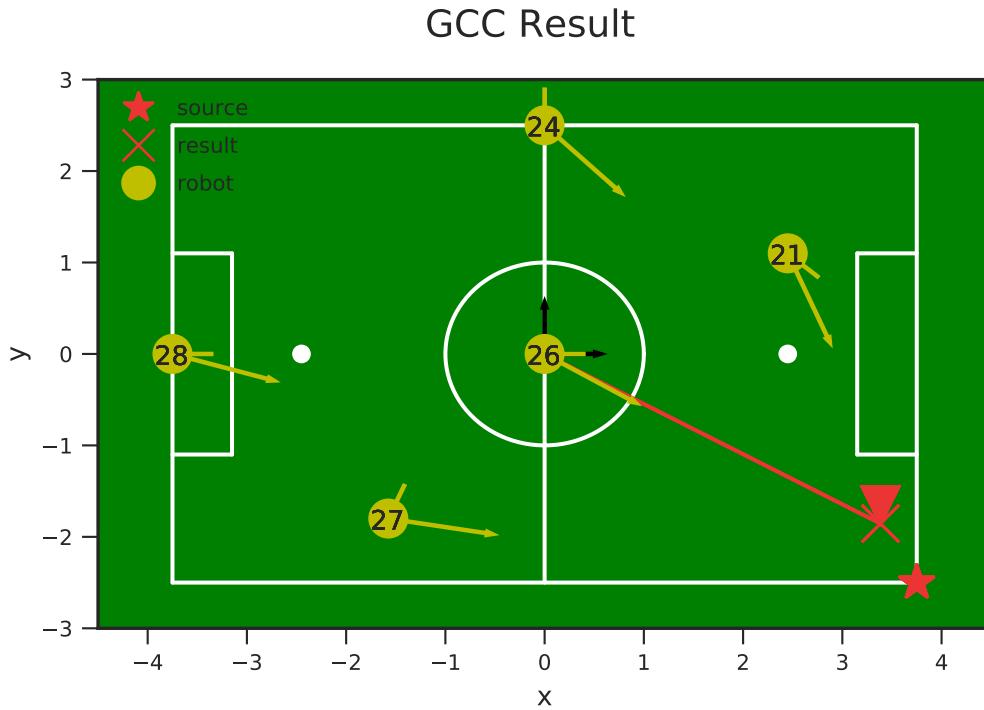


Figure 4.10: Team whistle localization result with GCC-PHAT method.

Measurement	Error x [m]	Error y [m]	Abs. Distance Error [m]	Angle Error
[0]	1,31	1,06	1,68	1,45°
[1]	0,13	0,06	0,15	1,57°
[2]	0,59	0,43	0,73	0,42°
[3]	0,54	0,47	0,72	9,09°
[4]	0,27	0,0	0,27	0,01°
[5]	0,15	0,14	0,21	3,18°
[6]	0,41	-0,02	0,41	2,67°
[7]	0,39	0,02	0,39	8,98°
[8]	1,84	-0,01	1,84	0,14°
[9]	0,58	0,52	0,78	9,89°
[10]	0,03	-0,0	0,03	0,0°

Table 4.11: Whistle localization results for all measurements in section 4.3.1 with GCC-PHAT method.

section 2.3 prove to be true.

Further information about the origin of the error can be obtained by looking at the single robot results of each measurement. Because the team filter is updated by the intersections of the single robot rays, the accuracy of the absolute position depends on the number of arising intersections. This will be discussed in section 4.3.5.

Measurement	Error x [m]	Error y [m]	Abs. Distance Error [m]	Angle Error
[0]	0,6	1,39	1,51	8,15°
[1]	-0,49	1,2	1,3	11,97°
[2]	2,32	2,11	3,13	18,28°
[3]	1,07	-0,96	1,44	3,71°
[4]	1,95	-0,09	1,95	10,44°
[5]	0,07	0,01	0,07	0,32°
[6]	0,4	-0,01	0,4	1,91°
[7]	1,06	-0,0	1,06	22,99°
[8]	1,24	-0,06	1,24	0,77°
[9]	-0,04	0,78	0,78	7,22°
[10]	0,03	-0,0	0,03	0,0°

Table 4.12: Whistle localization results for all measurements in section 4.3.1 with CC method.

4.3.4 Phase Method

For the results by phase method, the minimal frequency value parameter is set to 2700Hz. Just like for the other methods, the results of the measurements are shown in table 4.13. The RMSE according to the distance is in the same order of magnitude as the CC method with 1,33m. Having a much larger RMSE regarding the angle with 74,8°, especially these angular errors of measurements 6 and 10 catch one's eye. Both measurements are taken at the center point of the field. Because the absolute position is in an acceptable error range, these both angular results will be neglected from the error calculation. Thus, the angular RMSE of the phase method without measurements 6 and 10 is 11,69°.

Measurement	Error x [m]	Error y [m]	Abs. Distance Error [m]	Angle Error
[0]	-1,07	-0,98	1,45	4,13°
[1]	0,21	0,27	0,34	4,27°
[2]	0,22	1,39	1,41	16,25°
[3]	1,26	-0,02	1,26	11,19°
[4]	0,16	-0,04	0,17	0,99°
[5]	-0,42	0,19	0,47	5,46°
[6]	-0,32	0,08	0,33	166,25°
[7]	-0,28	1,76	1,78	20,82°
[8]	2,37	0,27	2,39	4,23°
[9]	2,04	0,29	2,06	24,89°
[10]	-0,32	0,0	0,32	180,0°

Table 4.13: Whistle localization results for all measurements in section 4.3.1 with phase method.

4.3.5 Single Robot Angle Error

To identify the difference between the methods, a closer look into the results of the single robots prior to the team filter is given. Figure 4.11 presents the RMSE considering the direction results of all five robots for each measurement in section 4.3.1. Additionally, the standard deviation of each measurement allows insight into the validity of the single robot results. As one can see, the standard deviation in degrees of the GCC method is significantly smaller in comparison to the phase method for most measurements. How this influences the reliability of the sound localization is subject of discussion in section 5.2.

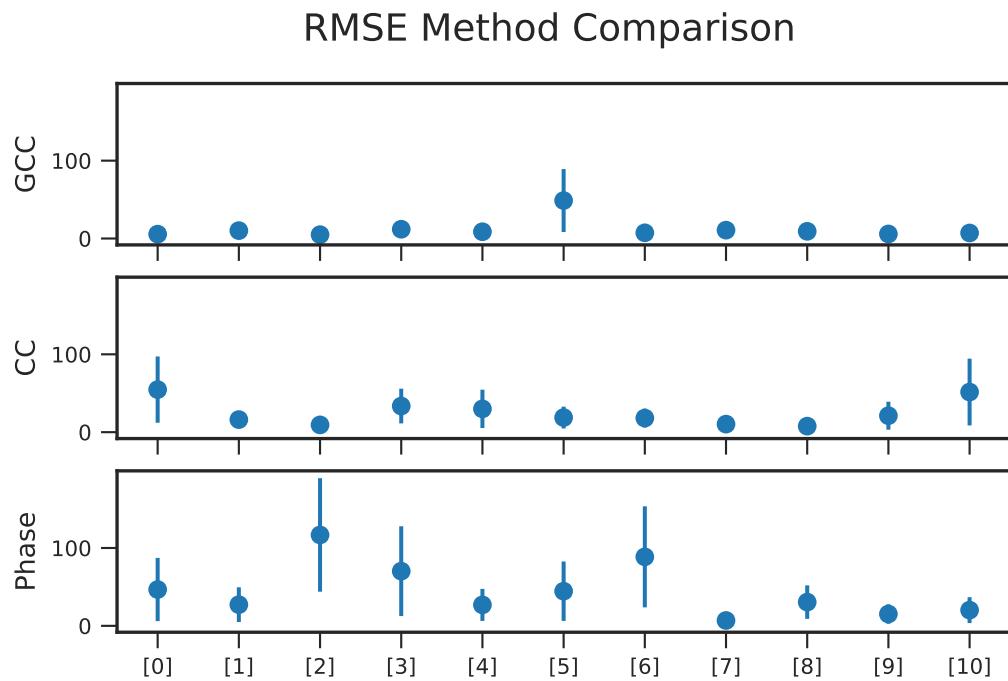


Figure 4.11: Angular RMSE and standard error of robot results per measurement of section 4.3.1.

4.4 Signal Start Detection

Previously, the importance of the signal start was underlined by the results of the previous chapters. Especially the result in section 3.3 has shown that in fact, the accuracy decreases with belated frame. Also according to the outcome of section 4.2.3, the GCC method performs best with frames near to the signal start.

In order to find a good solution for a highly reliable, accurate but computationally reasonable start detection, different approaches were tested profoundly. For high accuracy either the number of samples in a frame have to be small or the window must be shifted with small steps which both implicit a high number of executions. The existent whistle detection algorithm of section 3.1.2 is computationally intensive and causes some false positives for small frames as shown in . Therefore, it is looked for more simple algorithms that are not restricted to whistle sounds only.

Hereinafter, the characteristics of the methods are pointed out by analyzing the error between algorithmically determined start index and manually defined start index. For evaluation, the 11 measurements of on all robots section 4.3.1 are taken into consideration. By the frame size of the FFT being set to 256 samples for the correlation methods, a start index error of maximal 256 samples desired. Therefore, a start index detection result is regarded as failure for errors larger than 256 for the following sections. Two metrics are specified to express the degree of failure. One counts the number of large errors for each channel. In this case, 220 sample data for 11 measurements on five robots with four channels each exist. Another option is to count if any of the channels failed. For this case, 55 measurements are given.

4.4.1 Whistle Detection

First, the accuracy of the whistle detection in regard of the start index is evaluated. Here, the start index is defined as the first index where the whistle detection found a whistle. With a frame size of 1024, the whistle detection algorithm fails in every of the 55 measurements with at least one channel if only a error of maximal 256 samples is permitted. However, every error is in the range of 1024 samples which proves that the approximate start of the whistle sound is detected correctly. With a smaller frame size of 256, the error samples decrease but introduces false positive detections. With the results, one can say that the whistle detection is not sufficient for the start detection or is at least not reliable as stand-alone solution. Furthermore, this approach as it is limited to sounds in fixed and known frequency range.

4.4.2 ZCR

For the evaluation of the ZCR, the frame size is set to 256 samples and number of noise and signal frames are set to 10 frames. In 7% of the 220 channel sample data, the method fails with a significant error larger than 256 samples. In most cases, the method provides accurate results with small error. Taking all measurements of robot 26 as example, the RMSE amounts 54,23 samples.

However, there are cases where the algorithm fails. Looking at those cases it could be identified that the errors occur often due to incorrect assumptions that the signal is represented at the

end of the buffered samples. Some measurements prove, that this is not always the case as fig. 4.12 shows. Here, the data of the front right channel of robot 21 with the whistle source at position 5 in section 4.3.1 illustrates how the signal ended around 35000 samples already. If the start index is determined at the point where the ZCR falls below the threshold searching backwards as stated in section 2.7, the detection fails.

In most cases, only one channel of four output a erroneous result. Because the final start index on one robot is set equal for all channel, the failure can be compensated with a smart voting procedure.

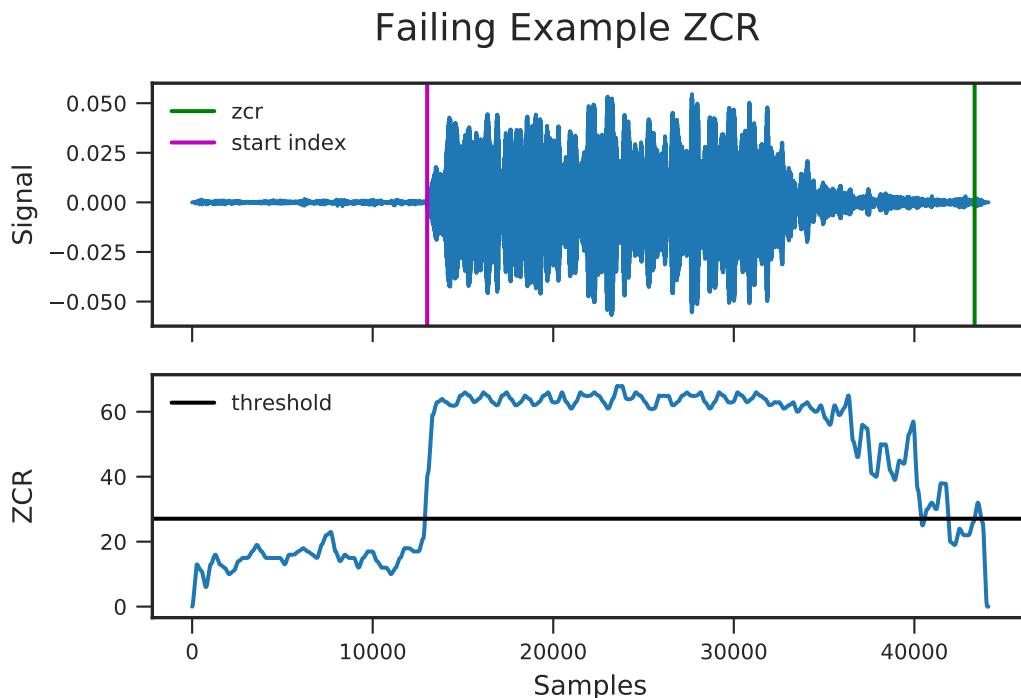


Figure 4.12: Channel 3 data from measurement 5 of section 4.3.1 for robot number 21.
A failing example for the start detection by ZCR is shown.

Another option exists by changing the process of finding the threshold excess onwards. In this case, the threshold is raised with a factor of 1,25. Results by this were poorer than the initially implemented manner with 15% failure rate. By adding the constraint that multiple samples must exceed the threshold successively, result can be slightly improved but implies higher computational effort.

Worth mentioning is the poor output of the ZCR method with signal that was cleaned with spectral subtraction previously. This surprising outcome is convenient for the overall task, because the start index result can be embed into the spectral subtraction, providing information for separating the noise and signal part.

4.4.3 Entropy

As discussed in section 2.7.2 the entropy quantifies the amount of chaos in a signal frame. Especially for signals to localize with unknown characteristics, this method can be useful because

no a priori knowledge is needed. For all measurements this method yield poorer results than the ZCR method with a failure rate of around 20%. Best results are achieved with a frame size of 512 samples and a difference step of 800 samples. However, for records with fading whistle the entropy method generates more reliable results than the ZCR method. Taking the same measurement as an example where the ZCR failed, fig. 4.13 shows how the algorithm detects the signal start correctly even though the whistle sound ended at around 35000 samples. In this measurement, the start index errors of all four channels were smaller than 40 samples.

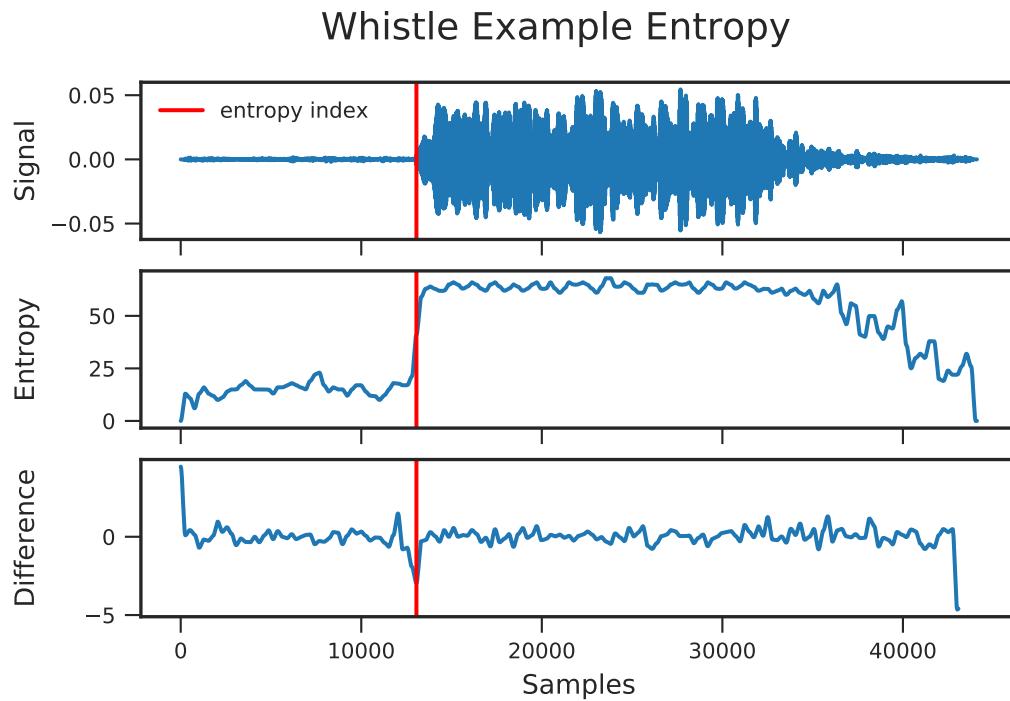


Figure 4.13: Exemplary result of start index detection by entropy where the ZCR method failed due to fading whistle at the data.