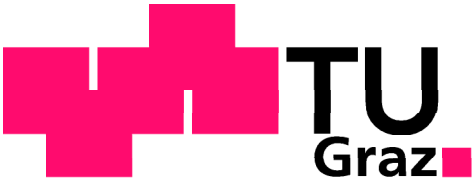

Graz University of Technology

Diploma Thesis:

Driving Direction Detection
with microphone arrays

Student
Peter Dollfuß, 0331114

Institute of Communication Networks and Satellite Communications
Head: Univ.-Prof. Dipl.-Ing. Dr.techn. Otto Koudelka

Assessor:
Univ.-Prof. Dipl.-Ing. Dr.techn. Otto Koudelka

Advisors:
Dipl.-Ing. Dr.techn. Franz Graf
DI Michael Stadtschnitzer

Date: May 27, 2010

Concerning my favourite hobby, the climbing I would say:

... a further cruxe of live is done [1] ...

But this wouldn't be possible without the help of great people.

First, special thanks goes to my parents whose made all possible, without their great psychological and financial support my studies wouldn't be that easy.

Thanks to Prof. Koudelka for the possibility to do this project in his company, the time in the joanneum was always a pleasure.

Thanks to my advisors Dr. Franz Graf and DI Michael Stadtschnitzer, for their time they spend with advising me on this great way.

For the great work done in preliminary and the support after handing over the first project parts, thanks to DI Harald Rainer, Martin Rohrmoser and Christoph Reitbauer.

Finally thanks to my sister Sabrina and Steve for their support in this time.

[1] cruxes means a difficult position in climbing (eine Schlüsselstelle), which is judging the whole tour

Zusammenfassung

In dieser Arbeit geht es um die Weiterentwicklung eines Fahrtrichtungserkennungssystems auf der Basis eines Beamformers (Mikrofonarray) mit nachgeschaltetem Fahrzeug bzw. Richtungsdetektor.

Um die theoretische Performance eines solchen Systems abschätzen zu können wurde als erster Teil der Arbeit eine Simulation eines Beamformers implementiert. Da die Software Matlab nur die Möglichkeit einer Offlinebearbeitung bietet und sich auch das Benutzerumfeld der aktuellen Version zu unflexibel darstellt wurde das Simulationtool in Labview realisiert. Labview ermöglicht die parallele (Verteilung auf mehrere Cores, CPUs möglich) Verarbeitung von Daten wodurch damit z.b eine direkte Analyse des Beamformerverhaltens bei sich änderenden Parametern möglich wird um z.B.: die Auswirkung von Sensorpositionierungsungenauigkeiten zu untersuchen.

Nach Übernahme der ersten Systementwicklung (Algorithmus, Array und Messdaten) wurden die Grenzen bzw. Probleme ermittelt, wobei sich speziell der Richtungsdetektor, bei hoher Verkehrsdichte, als zu instabil darstellte. Durch die Verwendung eines Algorithmus aus der Bildverarbeitung konnte zwar die Detektionshäufigkeit erhöht werden (und die Fehlerwahrscheinlichkeit gesenkt werden) jedoch auch nicht in dem gewünschten Ausmaß. Aufgrund einer Performanceanalyse des Beamformers und der akustischen Situation zeigte sich bald, dass auch bei Verwendung eines optimalen Auto/Richtungsdetektors die Probleme bei der Detektion durch die „schlechte“ räumliche Filterung verursacht wurden. Weitere Analyse zeigte die Notwendigkeit eines 2D Arrays anstatt des „2 Linien Arrays“ auf.

Schlußendlich resultiert aus der Arbeit ein System das bei hohem Verkehrsaufkommen und jeglichen Umweltbeeinflussungen (Nässe, Nebel etc.) überzeugt.

Abstract

This work is about the enhancement of a directional driving detection system. The system is based on a beamformer (microphone array) with two downstreamed binary hypothese detectors, a vehicle and a direction detector.

The first step in the work was the development of a beamformer simulation, to evaluate the theoretical performance of such systems especially in fields like on the freeway. Due to the fact that the software matlab provides just a offline calculation (and the Gui in the present version is not flexible enough) the Softwarepackage Labview of National Instruments was used. With this online system its possible to simulate the performance of the beamformer with varying parameters (for example erros in sensor positioning).

After handing over the first draft (algorithm, array and measured data) the adjustment of the program was done. Thus the problems and limits of the solution have been evaluated. Hereby it turned out that especially the direction detector is too unstable in high traffic.

By using an image processing algorithm, a much higher probabiltiy of detection and a lower probability of error could be achieved at the same time. However, based on a performance analysis of the system it could be presented that the most significant problems are caused by the spatial filter itself and could not be solved by enhancing the vehicle/direction - detector system. Everything pointed to the next step, the enhancement to a 2D system instead of the 2 particular line array.

Finally a system results which delivers stable results even in high traffic and with different enviromental conditions like fog, darkness and rain.

Contents

Chapter 1

Introduction

1.1 Motivation

> *... you are driving unsuspecting on the highway, suddenly there are car lights in your direction → a wrongway driver is on the way ...*

That sounds like a nightmare but in Austria it occurs on average 1.5 times a day that a car driver on the highway goes into the wrong direction. In the year 2009 there have been 390 wrong way drivers in sum (for comparison in germany 1500) and thus Austria is for the last years on the top of the european statistics (see the table below). The reasons for wrong way driving amongst are for the most part alcohol, suicide or adverse drive ups.

If it happens that such a wrong way driver causes an accident, it is fatal most of the time and unfortunately often deadly for the car occupants. On the account of this it is essential to detect such drivers as fast as possible, alert other drivers and the police (maybe automatic). Of course there is a good working society of drivers (Oe3ver etc.) whose raise alarm and alert the radio stations. But most of the time the particular passage of the highway is missed in the message or the radio is just off.

Hence the development of an automatic detection system for wrong way drivers, which is presented in this work is essentially for the safety of all motorists. The used system is an acoustic one and has the advantage of working with almost every enviromental conditions (fog, snow, rain, darkness etc.).

Table 1.1: Wrongway drivers of the last years in Austria [12]

Year	2004	2005	2006	2007	2008	2009
Driver	552	528	495	522	504	489

1.2 Preliminary work

The basic idea and the first drafts of the system are based on the great work of the Joanneum team.

Overview of the done work:

- Creating the basic idea,
- Building up of the hardware for the first system,
- Programming the first software solution,
- Setting up and programming the measuring hardware,
- Annotation of the collected data,
- and much more ...

The work that has been done in the front end was great and can't be covered in a few list items. Especially the work besides the direct development shouldn't be underrated there had been so much organisation, evaluation and annotation stuff which is really time-consuming, but essential and nessecary for the success of the project.

In the chapter 5.1.3 there will be a short description of the „first"software code which is important to understand the next steps. But if the more particular work is of note, I am free to refer to the works.

1.3 The basic idea

The general task is the development of a system for the acoustical detection of wrong way drivers. In fact, the exact realisation is based on a beamformer system (microphone array) in combination with 2 detector stages.

The beamformer [1], what can be seen as a spatial filter, samples the road surface and measures thus the sound pressure in special directions (which has been defined before in terms of defining a measuring grid on the road) or just generates an „acoustic video". A crossing motor vehicle or even a couple of vehicles (which is the normal case on freeways) will produce moving peaks on the acoustic video. The displacement of these peaks is due to the driving direction and has only to be detected by the system.

As can be seen in the figure 1 representing the basic idea, there is a box including the microphone array mounted on the right side of the traffic-gantry looking onto the road surface (figure 3). For explaining purpose the road surface is indicated by a sampling net whereas each field is supplied with numbers. The two cones starting at the box

[1] The term beamformer will be explained later in 1.4

just symbolises 2 beamformer main lobes steered on two grid-fields [2](in our case 10 and 11). With this fact it becomes clear if the blue car is crossing the grid area it will be detected first with the beam steered on the field 10 and then on field 11 or the other way around. With the knowledge of the detection direction and the temporal delay between the detections the direction, even the speed of the vehicle can be determined.

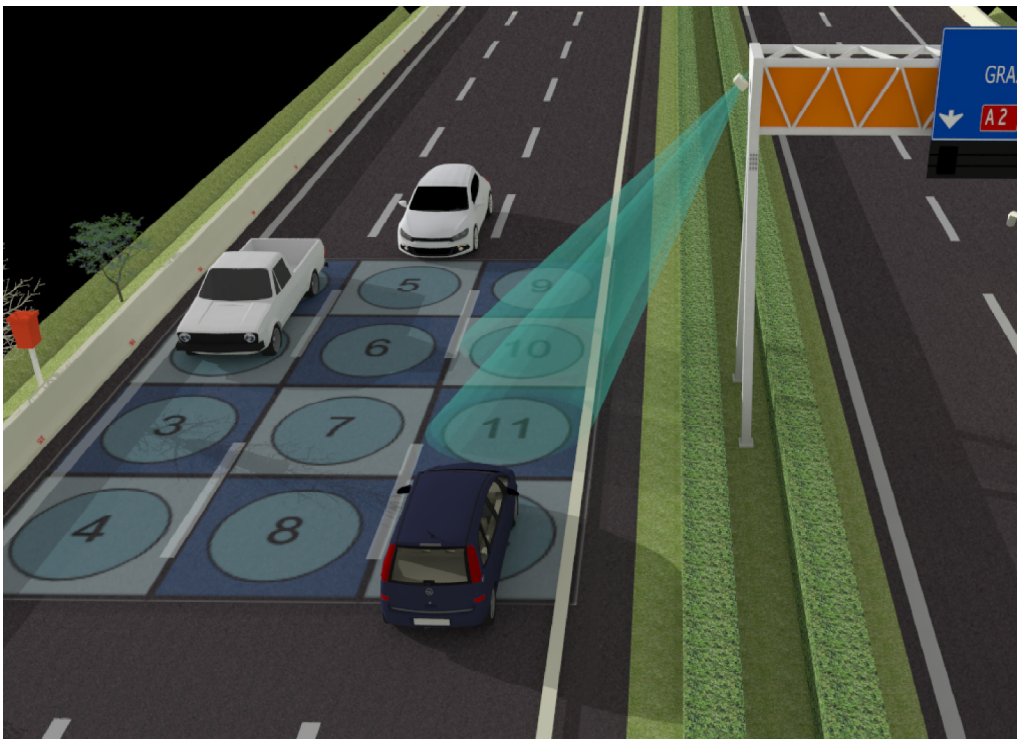

Figure 1: *The sketch illustrates the system for the explanation of the basic idea*

As identified above the whole system can be broken into 2 fundamental parts, whose work almost independent (see figure 2):

1. The spatial filter / Beamformer
2. The detector unit, which is composed by
    - Car detector
    - Direction / speed detector

[2]For explaining purpose the grid is coarse and the beampattern just represented by the mainlobe

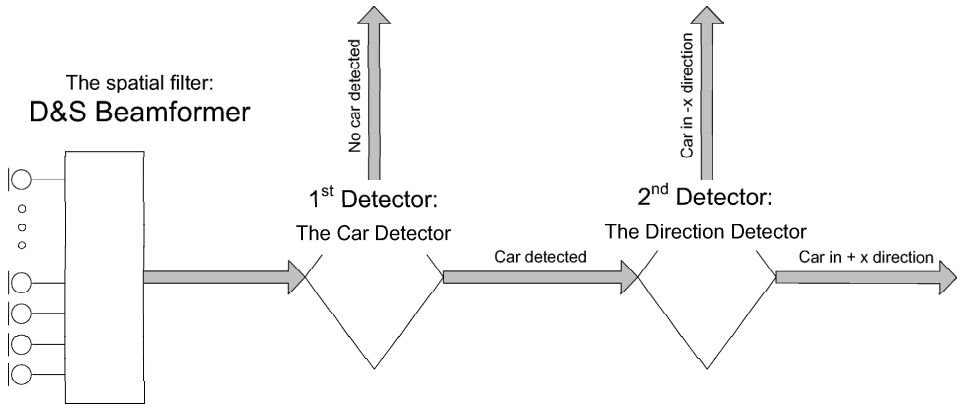

Figure 2: *The figure shows the general build up of the system by a block diagram*

Figure 3: *The picture illustrates the real system mounted on a highway gantry*

1.4 The basic (performance) quantities

Concerning the next chapters it is essential to introduce the spatial domain and to present the basic quantities describing beamforming system.

1.4.1 The signal model

For explaining purpose the simplest beamfomer, a narrowband sum and delay beamfomer is assumed. The description starts therefore with a common signal model in acoustics, a solution of the linear wave equation which can result, depending on the special boundary conditions in a plane or a spherical wave equation. With beamforming a spherical solution is choosen to describe the wavefield of the source:

$$s(\vec{p},t) = S\, e^{\overbrace{j(w_0\, t}^{temporal\ term} - \overbrace{\vec{k_0}\, \vec{p})}^{spatial\ term}} \tag{1.1}$$

By assuming to be in the farfield of the source (farfield beamformer) two approximations can be done:

1. Farfield approximation

Due to the infinite distance between the source and the array we can neglect the curvature of the spherical wave. This leads to the point that the spherical wave can be assumed as a plane wave [1].

2. Farfield Approximation

Due to the infinite distance between the source and the array there is no big impact on the signal strength → the microphones having roughly the same distance to the source.

On a more precise look the assumptions apply to the case where **the dimensions of the array are much smaller then the wavelength of the source** (seen from the beamformer).

Further approximations

Furthermore the source is assumed to be in a **free field** what means no acoustic restriction and just the direct propagation path. But as can be seen later, there is of course restriction by a surface (see later in section 3.1) and thus depending on the surface absorption a first reflection additional to the direct path.[1]. Additionaly it has to be said that the model is based on a linear system but with a more precise look it can be seen

[1] In fact this is a general assumption due to the shape of the wavefronts but the amplitude is still decreasing with $\frac{1}{r}$ and not constant like with real plane waves

[1] In special acoustic situations → for example with soundproof walls or a bridge above the street a more complex impulse response will occur and thus a multipath propagation model should be assumed

that with moving sources the „doppler shift“occurs and leads to non linearities. But since the car is just tracked in a small area, the situation can be seen as quasilinear [2].

1.4.2 Spatial and temporal sampling

By using a (discrete) microphone array the wavefield is sampled in space. Additionally for further processing the particular microphone signals are sampled in time by the A/D converters as well. The sampling in time and in space can be mathematical modeled by multiplying the continuous signal with a pulse train [6],[3]:

$$x_m = s(\vec{p}, t)\, \delta(\|\vec{p} - \vec{p_m}\|_2)\, \delta(t - k\, T_s) \tag{1.2}$$

The microphone positioning vector p_m can be expressed with the spatial sampling periode vector $\vec{d}$ as well:

$$\vec{p_m} = \vec{p_0} + m\, \vec{d_m} \tag{1.3}$$

x_m	...	output signal of the m-th microphone channel
$s(\vec{p}, t)$	...	the source wavefield
$\vec{p}$	...	position vector
$\vec{p_m}$	...	microphone position vector
$\vec{d_m}$	...	spatial sampling vector
T_s	...	temporal sampling periode

With further considerations of the mathematical sampling modelation (in time and in space) it turns out that from spatial sampling the same effect occurs like in the temporal case → the spectra will be periodic continued. The following table should illustrate the relations through the domains with their parameters:

time domain →	**frequency domain**
time t	circular frequency ω
spatial domain →	**k-space or directional domain**
position (x,y)	circular wavenumber vector $\vec{k}$

In fact the k or the circular wavenumber can be expressed by two angles ϕ_1 and ϕ_2:

$$\vec{k} = \frac{2\pi}{\lambda}\vec{k_e} = \frac{2\pi}{\lambda}\sin(\phi_1)\cos(\phi_2) \tag{1.4}$$

[2] like in speech communication the quasistationarity

Again with sampling in space the spatial spectra becomes periodic continued

To prove this statement the spatial response of a gridarray is calculated and plotted in matlab (figure 4),therefore in consequence the script cut-out is denoted:

```
%-----------------------------------------------------------------
% defining the spatial filter:
H=ones(4,4) % 16 microphones with rectangular array design
% defining the zero padding in x and y direction
Nx = 32; Ny = 32;
% Defining the relation r=lambda/d;
  r=0.5;
% Calculation and plot of the response
  freqz2(H,Nx,Ny,r)
%-----------------------------------------------------------------
```

The function freqz2 (in fact realized by a 2D FFT) gives the possibility to zero padding. The frequency and the distance between the microphones is defined by the value of r, the relation between the wavelength λ and the distance d. In our case r is initialized by 0.5 what means the wave is 2 times sampled.

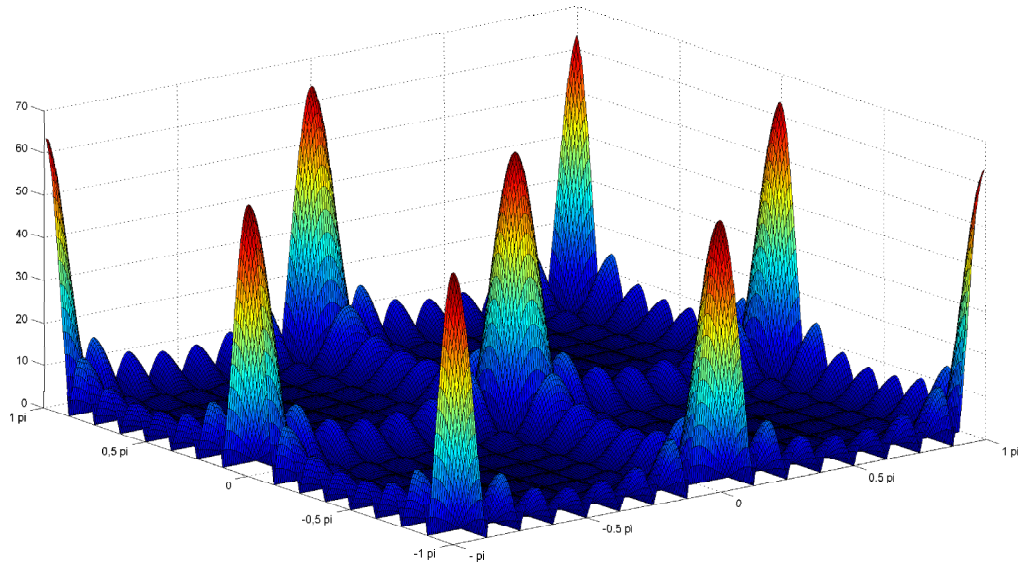

Figure 4: *The plot of the spatial response due to a grid array*

The figure 4 shows the result of the calculations, the spatial response (this is indeed alread the beampattern), whereas the periodic continuation of the area around the main lobe becomes obvious. But its really important to note that the area around the main lobe is limited by a circle.

To get back again to the spatial sampling, it have to be said that especially in space it is already common to sample irregular, what means the distances between the microphones are not equal. In other words it is often helpfull, depending on the sourcesignal to use different array designs [3].

1.4.3 Spatial aliasing

As mentioned before in equation 1.2 the wavefield is sampled in space with $\delta(\|\vec{p} - \vec{p_m}\|_2)$ whereas $\vec{p_m} = \vec{p_0} + m\,\vec{d_m}$.

d_m means the spatial sampling periode vector like in the temporal domain the sampling periode T_s. With this analogue the spatial aliasing theorem can be derived:

In the time domain the nyquist theorem is defined as follows:

$$T_s \leq \frac{T}{2} \rightarrow f \leq \frac{f_s}{2} \tag{1.5}$$

With T_s similar to $\vec{d_m}$ and T is equivalent to $\vec{\lambda}$ the spatial aliasing theorem can be derived:

$$2\,\vec{d_m} \leq \vec{\lambda} = \frac{2\pi}{\vec{k}} \rightarrow d_m \leq \frac{\lambda}{2\cos(\phi)} \tag{1.6}$$

With this equation it becomes obviouse that with every particular distance between the microphones the theorem has to be valid, and additional its depending on the direction of the incident wave. Later in the beampattern calcuation section the effect of spatial aliasing will be presented with some examples.

1.4.4 The spatial response, Beampattern

In section 1.3 it has been stated that the beamformer can be seen as a spatial filter. To prove this the beampattern, which is actually the spatial response of the beamformer is calculated.

Point of orgin for our calculations is the equation for the particular microphone signals (spatial sampled signals) [6],[3]:

$$x_m(t) = \alpha_m\, s_m(t - t_0 - \Delta_{m,r}(\phi)) \tag{1.7}$$

[3] In temporal domain just a few systems use the irregular sampling to reduce or even avoid aliasing effects actual the energy in the spectra is leaked etc. but with the disadvange of a more complex sampling system (the point of time for every particuar sampling value has to be known -> the sampling rate is not enough) and more noise

In this equation the m indicates the m-th microphone signal, t_0 the time from the source to the reference sensor and the relativ delay $\Delta_{r,m}$. The Delay therefore is depending on the incident angles of the wave. Finally a damping factor α_m is introduced:

With the farfield approximations equation 1.2 can be reduced to:

$$x_m(t) = s_m(t - \Delta_{m,r}(\phi)) \tag{1.8}$$

By applying the first processing step of the sum and delay beamformer the signal is aligned:

$$x_{m,d}(t) = x_m(t + \Delta(\psi)) = s_m(t - \Delta_{r,m}(\phi) + \Delta_{r,m}(\psi)) \tag{1.9}$$

It is of course desired to match with the beamformer delay the real occured delay, thus with farfield assumptions the delay calculations can be done by:

$$\Delta_{r,m}(\psi) = \frac{m\ d\ cos(\psi)}{c} \tag{1.10}$$

Another possibility (which is choosen in the simulation) calculates the delay by the vector difference between the source position vector $\vec{Sp}$ and the microphone position vector $\vec{Mp}$ (the index $_2$ indicates the euclidian vector norm):

$$\Delta_r, m(\vec{Sp}) = \left\|\vec{Sp} - \vec{Mp}_m\right\|_2 - \left\|\vec{S_p}\right\|_2 \tag{1.11}$$

It can be shown that the vector delay calculation converges with large distance to the common far field solution.

The second processing step due to sum and delay beamforming is the summation of the channels (additionally a weighting value is adverted → explanation later):

$$y_{ds}(t) = \frac{1}{M} \sum_{m=0}^{M-1} \omega_m\ x_{m,d}(t) \tag{1.12}$$

By inserting the expression (1.8) it becomes obvious that the summed signal is dependent on the time, the indicent angle of the wave ϕ and the steering angle ψ:

$$y_{ds}(t, \phi, \psi) = \frac{1}{M} \sum_{m=0}^{M-1} \omega_m s_m(t - \Delta_{r,m}(\phi) + \Delta_{r,m}(\psi)) \tag{1.13}$$

By using the frequency version equation 1.13 results in:

$$y_{ds}(\omega, \phi, \psi) = \frac{1}{M} \sum_{m=0}^{M-1} \omega_m s_m(\omega) e^{-j\omega\Delta_{r,m}(\phi)} e^{j\omega\Delta_{r,m}(\psi)} \tag{1.14}$$

By using the farfield delay calculation 1.10:

$$y_{ds}(\omega, \phi, \psi) = \frac{1}{M} \sum_{m=0}^{M-1} \omega_m s_m(\omega) e^{-j\omega\, m\, d\, \frac{cos(\phi-\psi)}{c}} \tag{1.15}$$

By assuming constant source power $s(\omega) = s_0$ and $\omega_m = 1$ (weighting value):

$$y_{ds}(\omega, \phi, \psi) = s_0 \frac{1}{M} \sum_{m=0}^{M-1} e^{-j\omega\, m\, d\, \frac{cos(\phi-\psi)}{c}} = s_0 \frac{1}{M} \sum_{m=0}^{M-1} e^{-j\frac{2\pi}{\frac{\lambda}{d}}\, m\, cos(\phi-\psi)} \tag{1.16}$$

By considering equation 1.16 the form the cause why beamformer are often called spatial filters becomes obvious. The form of the equation is similar to the form of a temporal DFT (except from the cos part) but of course with different parameters. By assuming a gridarray the angles elevation and azimuth are used, what leads to the before mentioned circle limited area and to axis distortions.

In the next step the beampattern is calculated [3]:

$$B_p(\omega, \phi, \psi) = \left| \frac{sin(M\, \pi\, \frac{d}{\lambda}\, cos(\psi - \phi))}{M\, sin(\pi\, \frac{d}{\lambda}\, cos(\psi - \phi))} \right| \tag{1.17}$$

1.4.5 Deduced quantities

Because the Beampattern presents the spatial response of the filter, its essential to evaluate the behaviour of such a system. With the Beampattern almost all the quantities of the beamformer system can be calculated. Just some of them are:

- The resolution,
- The maximum sidelobe level,
- The position of the sidelobes and much more ...

The mainlobe width

The mainlobe width is close connected to the resolution of the beamformer and because of this a important quantity. The calculation of the mainlobe width can be done easily by assuming a steering angle of 90° (thus in our case steered straigt forward) and by finding the second null in the Beampattern [4]:

[4] Sometimes the Beampattern is just defined by the distance between the -6 dB decreasing of the mainlobe

$$\sin\left(\pi\, N\, \frac{d}{\lambda}\, \cos(\psi)\right) \stackrel{!}{=} 0 \;\rightarrow\; \psi = n\, \cos^{-1}\left(\frac{\lambda}{M\, d}\right) \tag{1.18}$$

Equation 1.18 results in a general description of the positions whose leading to zeros in the beampattern (the integer n terms the number of the zero). In the case of calculating the mainlobewidth n should be one and the result ψ is taken two times.

The angular resolution

As mentioned above, the resolution of the beamformer is directly connected to the mainlobe width. For this reason a look on the definition of the resolution by the rayleigh criterion is done:

> *The rayleigh criterion defines that 2 waves incident from directions close to each other can be seperated if the peak of the 1st incident wave falls into the 1st zero of the second wave.*

In other words: the angular distance between 2 sources have to be at least the half mainlobe width. Till then its possible to regognize 2 seperate sources instead of one. Until now the beamwidth and the resolution has been considerd just angular, but for real solutions there have to be a look on this quantities on a focus plane as well. For the calculation of Bm (the mainlobe width on the focus plane) just trigonometric functions are used:

$$tan(\psi) = \frac{\frac{Bm_{axis}}{2}}{z} \rightarrow Bm_{axis} = 2\, z\, tan(\psi) \tag{1.19}$$

Cause Bm is additionally dependent on the steering angle ϕ and the term is more complex the equation relating the on axis and the off axis resolution has been inccured from [2]:

$$B_m(\phi) = \frac{1}{\cos^3(\phi)}\, Bm_{axis} \tag{1.20}$$

By observing the values in the plot, it can be seen that already with $\pm$ the beamwidth is 2 times larger than on axis. For that reason often the beamformer system doesn't use steering angles above $\pm$ 30°.

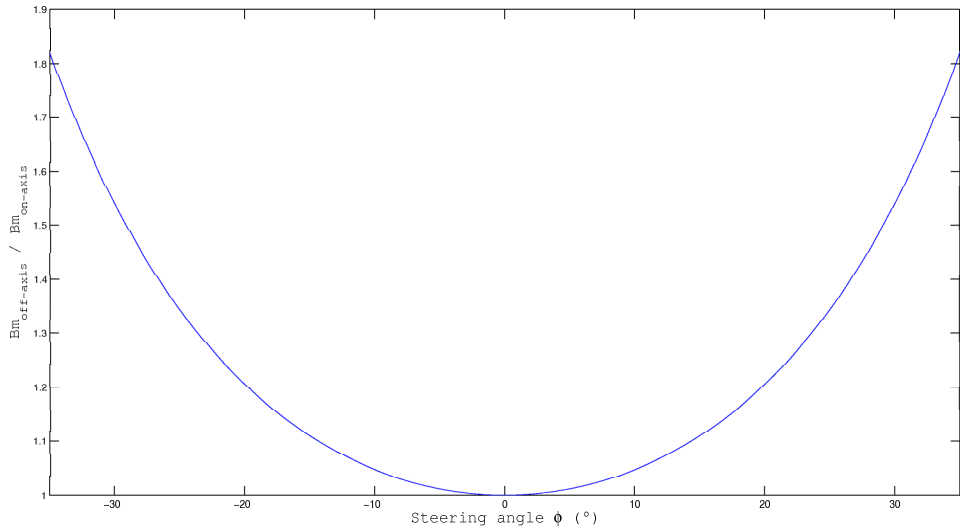

Figure 5: *The plot shows the relation between the on axis resolution and the off axis resolution*

The maximum sidelobe level

Since sidelobes can be seen as spatial leakage (due to the limited aperture) it is not possible to distinguish between real signals from the sidelobe direction or just leakage. Thus it becomes clear that it is desired to get low sidelobe in relation to the mainlobe(like in the temporal domain the sidelobe levels can be reduced by window function but with worse mainlobe width). The quantity which describes this ability is called Maximum Sidelobe Level or MSL:

$$MSL = 20\ \log\ \frac{A_{mainlobe}}{A_{sidelobe}} \tag{1.21}$$

Before calculating the Maximum Sidelobe Level the calculation of the sidelobeposition must be done (finding the 2nd maxima in the beampattern):

$$\psi_{sidelobe} = \left(2\,n + 1\,\frac{\lambda}{2\,M\,d}\right) \tag{1.22}$$

With $n = 1$ (1st sidelobe) and the aperture function the level of the 1st sidelobe becomes:

$$A_{sidelobe} = \frac{1}{M\ \sin(\frac{3\,\pi}{2\,M})} \tag{1.23}$$

From equation 1.17 it results that the position and the level of the sidelobes are dependent on just the amount of the microphones. Furthermore with increasing number of microphones the sidelobe level will decrease and more sidelobes in general will occure.

1.4.6 Different approaches

Beside standard quantities, whose orginate to the filter theory there are some different approaches [7]:

The ambient noise gain

For explaining a diffuse ambient noise field has to be assumed: an ambient noise field can be imagined by an infinite amount of sources distributed on a sphere around the array. Due to the directivty of the array (apart from low frequencies where the beampattern is almost spherical) just the noise in the direction of the lobes is accounted → thus the array output results in lower noise compared to a single microphone with its spherical characteristic.

$$r = \frac{d}{\lambda} \tag{1.24}$$

The used term in equation 1.24 describes the relation between the microphone distance and the wavelength, thus can be seen as a kind of normalisation.

$$G_a = \frac{1}{N_\phi} \int_{+\pi}^{-\pi} B_p(\frac{d}{\lambda}, \phi) d\psi = \frac{1}{N_\phi} \int_{+\pi}^{-\pi} \left| \frac{sin(M\ \pi\ \frac{d}{\lambda}\ cos(\psi)}{M\ sin(\pi\ \frac{d}{\lambda}\ cos(\psi))} \right| d\psi \tag{1.25}$$

The equation 1.25 is of course in the normal case a volume integral (a 3 times ring integral) divided by the work volume.

The instrumental noise gain

The white noise gain (WNG) is defined as the ability to suppress spatially uncorrelated noise (e.g. internal noise of the microphones):

$$G_n = \sqrt{\sum_{m=1}^{M} (\alpha_m)^2} \tag{1.26}$$

As can be seen the noise gain is just dependent on the α_m, the microphone channel weighting.

1.4.7 Summary

The summary can be carried out by using an calculation example and by discussing the results.

Calculation-example

Array design: linear array with equi-spaced microphone distance

M=4	...	the amount of microphones
ϕ=0°	...	the steering angle relative to 90°
r=0.5,0.7,1	...	see equation 1.24

Spatial aliasing:

$$\frac{d}{\lambda} \leq \frac{1}{2\ \cos(\psi)} = \frac{1}{2} \tag{1.27}$$

With the worst case, a wave incident against the microphone position vector ($\psi = 0°$) the equation 1.6 results in $\frac{d}{\lambda} \leq 0.5$. In the calculation example the used values for r are 0.5, 0.7 and 1, thus with our aliasing calculation the nyquist theorem is broken by 0.7 and 1. This fact can be seen in figure 6, whereas the polar plot is more helpfull. The aliasing effect can be seen obvious in the red polar plot which shows the beampattern for r=1 thus a spatial sampling period equal to the wavelength. As with the well known effect in the frequency domain aliasing leads to a false-interpretation of frequency or in space in a false interpretation of direction. The figure 6 or 7 shows the false main lobe, which is actually a mirror-mainlobe moved into the used spatial area (The mirror mainlobes are also called grating lobes). With the value r=0.7 (blue plot) the aliasing effect does not become obvious but by considering the sidelobewidth (which should stay constant) the effect already can be seen.

The angluar beamwidth: Bw_{ang} for r=0.5:

$$\psi = cos^{-1}\left(\frac{\lambda}{M\ d}\right) = cos^{-1}\left(\frac{1}{4\ 0.5}\right) = 60° \rightarrow Bw_{ang} = 2\ (90° - 60°) = 60° \tag{1.28}$$

r	$B_{ang}(°)$
0.5	60°
0.7	41.8°
1	29°

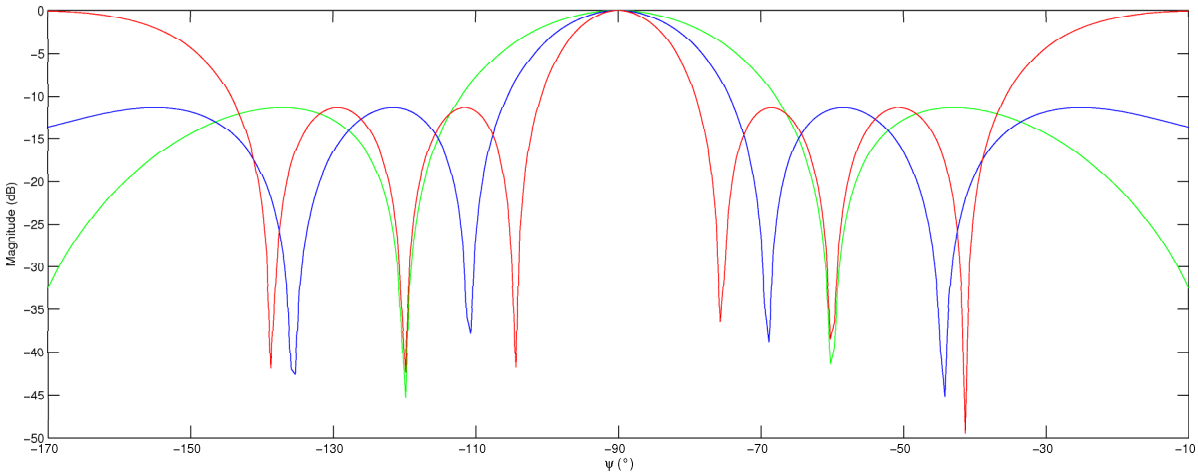

Figure 6: *The plot shows the beampattern on ψ with rectangular coordinates and in logarithmic scaling (r is 0.5 (green),0.7 (blue) and 1 (red))*

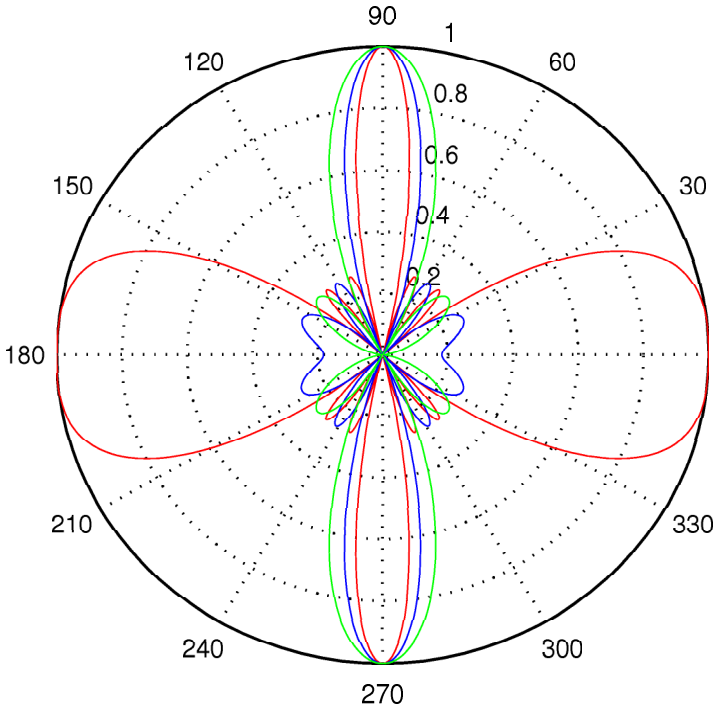

Figure 7: *The plot shows the beampattern on ψ with a polarplot and linear scaling (r is 0.5 (green),0.7 (blue) and 1 (red))*

By regarding the results it turns out that the (angluar) beamwidth is dependent on:

- the wavelength (frequency) $\rightarrow$ with increasing frequency the mainlobe becomes smaller.
- the aperture $M \cdot d$ with increasing aperture the mainlobe becomes also smaller.

The figure 8 shows obvious the dependency of the beamwidth, the sidelobes etc. on the frequency, whereas the microphone distance is choosen by 10 cm.

Again the spatial aliasing comes into role, because with the aperture-increase the amount of sensors (microphones) have to be increased too, to ensure the maximum distance between the microphones. A solution which deals with this problem is presented in [13] → the „nested"arrays. Additionally it has to be mentioned that with real beampatterns the zerocrossing (which is assumed for the calculation) will never be reached → thus an approximation has to be used.

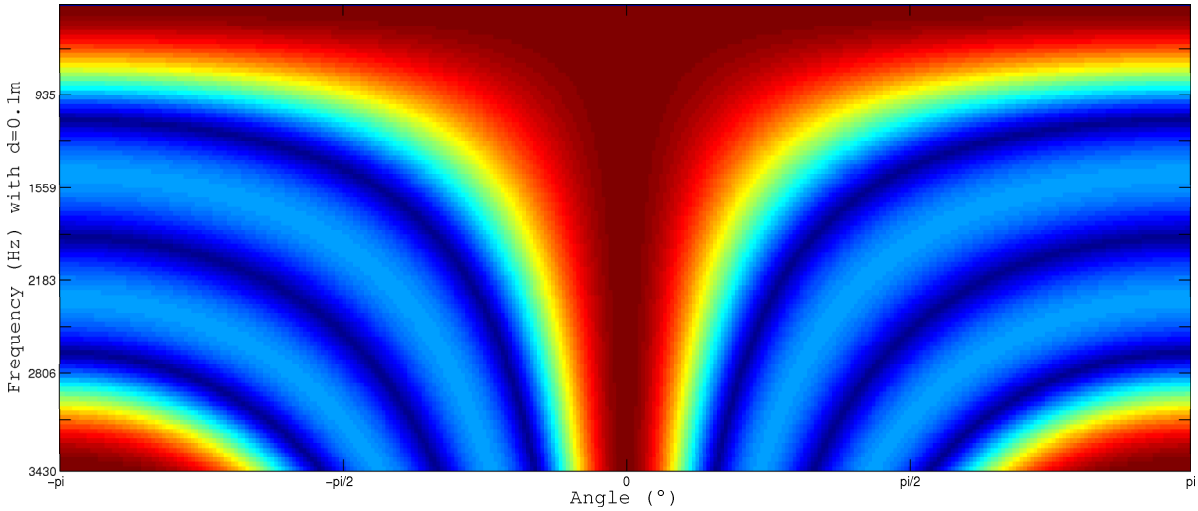

Figure 8: *The plots shows the frequency dependency of the beampattern for d=0.1m*

The sidelobe positions with r=0.5

$$\psi_{sidelobe} = \cos^{-1}\left((2\,n+1)\,\frac{\lambda}{2\,M\,d}\right) = \cos^{-1}\left((2\,n+1)\,\frac{1}{2\,4\,0.5}\right) = 48.6° \tag{1.29}$$

r	1^{st} sidelobe ψ and $n=1(°)$	2^{nd} sidelobe ψ and $n=2(°)$
0.5	48.6°	complex
0.7	32.4°	63.2°
1	22°	38.6°

With the sidelobe postion or in general with sidelobe results it can be seen that with increasing r, the sidelobes moving towards the mainlobe. The amount of sidelobes is interesting too because it results that $M+1$ sidelobes will occur but just with r=0.5 and over the whole area, what means from 0 to 2π.

The maximum sidelobe level

$$A_{sidelobe} = \frac{1}{M\ \sin(\frac{3\pi}{2\,M})} = \frac{1}{4\ \sin(\frac{3\pi}{2\,4})} = 0.2706 \rightarrow -11.3\,dB \tag{1.30}$$

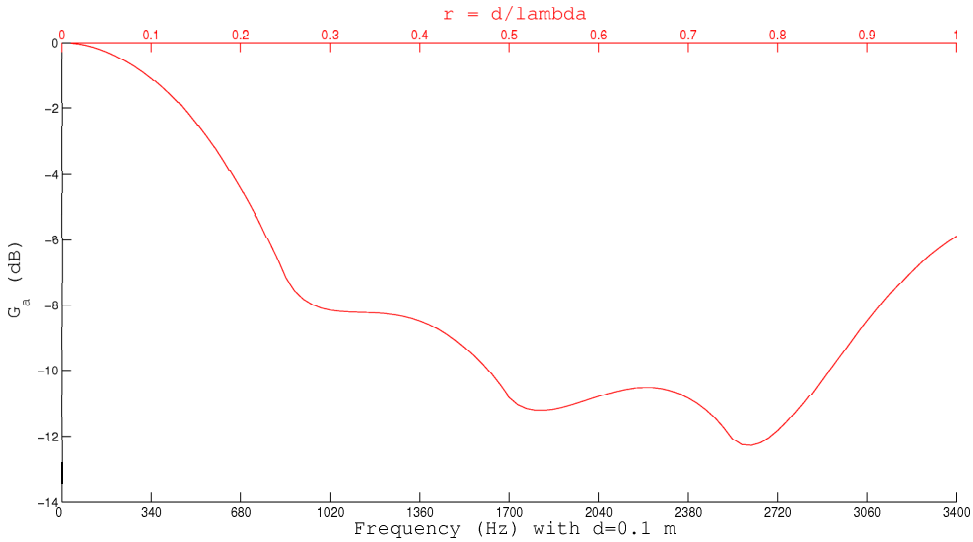

Figure 9: *The plot shows the dependency of the ambient noise gain on the frequency (or d)*

By calculating the sidelobe level it can be seen again that this quantity is for regular arrays just dependent on the amount of microphones.

The ambient noise gain
With a line array and its, „elevation independent"beampattern it is possible to calculate the ambient noise gain [7] with just the surface under the curve (steering angle $\phi = 0$ and N_ϕ is the length of ϕ):

$$G_a = \frac{1}{N_\phi} \int_{+\pi}^{\pi} B_p(\frac{d}{\lambda}, \phi) = \frac{1}{N_\phi} \int_{+\pi}^{\pi} \left| \frac{sin(M\ \pi\ \frac{d}{\lambda}\ cos(\psi)}{M\ sin(\pi\ \frac{d}{\lambda}\ cos(\psi))} \right| d\psi \tag{1.31}$$

With this quantity further values for r are choosen and the resulting function is plotted. The plot is viewed in figure 9 and shows a function whith 3 local minima, whereas the lowest minima is in the forbidden area ($r \approx 0.8$). It turns out that the ambient noise gain becomse best with 2 times sampling the wavelength. With real signals and arrays it is of course just valid for only one frequency but can be used to design arrays for narrowband signal sources.

The instrumental noise gain
The instrumental noise gain is just depending on the weighting values of the microphone channels, thus delivers with a spatial rectangular window (which shows the smallest mainlobewidth) the best result ($m = 4$):

$$G_n = \sqrt{\sum_{m=1}^{M} (\alpha_m)^2} = \sqrt{4} \rightarrow 6dB \tag{1.32}$$

Finally it has to be said that with irregular arrays the derivation of the quantities like mainlobe, sidelobes etc. is not really easy and can only be estimated by simulation. The only quantities whose are really helpful with irregular arrays too are the two gains, once the ability to supress diffuse noise from different directions and on the other hand supressing the internal (microphone, preamp noise) by averaging through the channels. But the instrumental noise gain assumes of course non correlated noise through the channels and delivers worse results with correlated ones (like noise from a specific point).

Furthermore at this point the beamformer with its beampattern can be seen as well like a rectangular window, which cuts out an area of the infinite wavefront and thus the resulting beampattern reminds to the deflection effect.

Chapter 2

Beamformer - Simulation

As mentionend above, most of the quantities of irregular beamformers can just be simulated and thus theorectical estimated. On that reason a simulation was developed to analyse proper and new array design. The simulation was planned in matlab and concerning the unflexibly Gui and the non-online calculation realised in Labview.

2.1 NI Labview 8.2

Labview is a visual based development and measurement environment which convinced by easy and fast handling and thus intergration of measurement tools.

Basically Labview is organized like a normal measurement instrument with a user interface and its function or realisation on the hardware layer (the files are calles vi's → virtual instrument). In Labview the user interface is called front panel and the functional layer is the block diagram, where the signal processing is realised.

2.2 Desired requirements

Before developing the simulation, the requirements have been fixed. Like with the most simulation programs, the general requirements are fast calculation and an user optimized (user friendly) system.

The functional requirements:

- Flexible array configuration (array design, amount of microphones and weighting of the channels)
  - Array design
  - Amount of microphones
  - Weighting of the channels
- Calculation of the beampattern for a sphere and specific coordinates
- Calculation of beamformer-quantities

2.3 Principle of simulation

The simulation is based on the narrowband realization of a sum and delay beamformer. Point of orgin is a basic technic of signal processing, which describes filters by its eigenfunctions. In the case of a spatial filter complex exponential functions are used:

$$s(t) = \cos(\omega t) + j\ \sin(\omega t) = e^{\,j\omega t} \tag{2.1}$$

The beampattern relates to the amplitude of the spatial filter and can be derived by calculating the filter coefficients of the beamformer, thus the steeringdelays. Additionally the source or the inputsignal of the spatial filter has to be considered, which is easily done by accounting the propagation delay of the source point to the sensors t_{ms} (calculation see section 1.4).

Again by processing the sum and delay beamformer on the complex exponential signal the beamformer output results in:

$$y(t, \vec{\kappa}) = \sum_{m=1}^{M} \left(\cos_m(\omega t_{ms}) + j\ \sin_m(\omega t_{ms})\right) e^{-j\,\omega \Delta_m(\vec{\kappa})} \tag{2.2}$$

By assuming a steering angle $\vec{\kappa} = 0$ the equation 2.2 becomes:

$$y(t, 0) = \sum_{m=1}^{M} e_m^{j\omega\, t_{ms}} e^{-j\,\omega \Delta_m(0)} = \sum_{m=1}^{M} e_m^{j\omega\, t_{ms}} \tag{2.3}$$

With 2.3 it becomes clear that with only one source just a single value is gotten. To calculate the whole beampattern „virtual“sources in every desired direction have to be defined. In the simulation this is done by defining sources liing on a sphere around the array. Furthermore to calculated the road-response the sources are defined on the road. The term response is choosen on purpose because by defining for example an arbitrary „calculation net“the calculations resulting more excactly in the response due to a special spatial signal and not in the beampattern.

2.4 The program

The program is sectioned into:

1. Stage: Value initialisation and precalculations (one times exexution)
2. Stage: Calculation of the beampattern, response and quantities
3. Stage: Plotting the results, saving the data

The basic programming structure in this project is a matlab code section for the initialisation, followed by a while loop which is executed as long as the stop button is not pushed. Cause the initialized data is used all the time it have to be save within a shift register structure. The particular parts will be commented or described in the following figures (thus in Labview).

2.4.1 The first stage: initialisation and precalculations

In this stage (see figure 10) the definition of the grids (sphere and surface), the frequency range, temperature etc. is done. The initialisition is done in a matlab script node which is useful in order to use the matlab function like sphere and meshgrid (this functions can be realized in Labview as well). After doing some coordinate transforms the script overhands the data to the subVI - Precalculations, which adapts and prepares the data for future calculations.

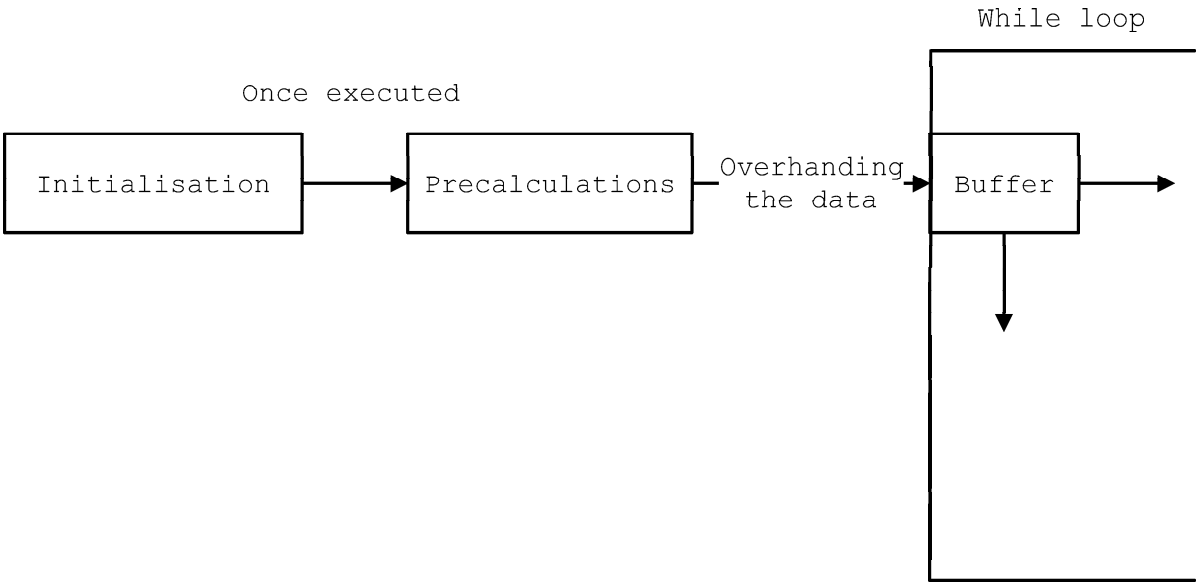

Figure 10: *The figure illustrates a block diagram of the* 1^{st} *stage*

2.4.2 The second stage: main calculations

In this stage (figure 11) the first step is the definitions of the microphone array, the array design (linear, rectangular, ...) and if desired the rotation along the x,y,z - axis is applied. With the microphone positions and the mesh, spheregrid the delays can be calculated in the next step (the calculation for the meshgrid and the sphere is done parallel). Finally the D&S stage will be executed, thus the eigensignal is delay and summed.

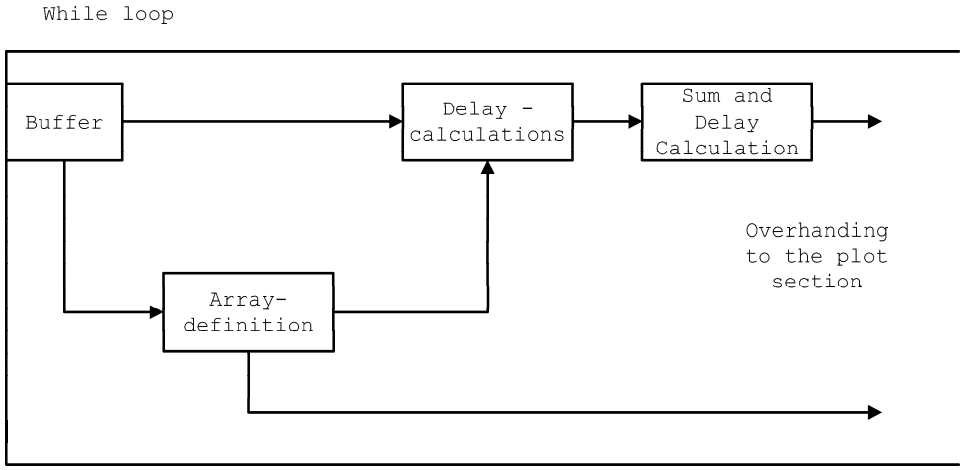

Figure 11: *The figure illustrates a block diagram of the* 2^{st} *stage*

2.4.3 The third stage: plotting and displaying the results

In this stage (figure 12) the results of the calculations are plotted:

- the microphone array with the sensor position and the steering vector
- the beampattern (the response on the sphere)
- the beamformer response on the surface

Additionally the calculated quantites, the ambient and the instrumental noise gain are display and the data can be saved for further processing (for example matlab).

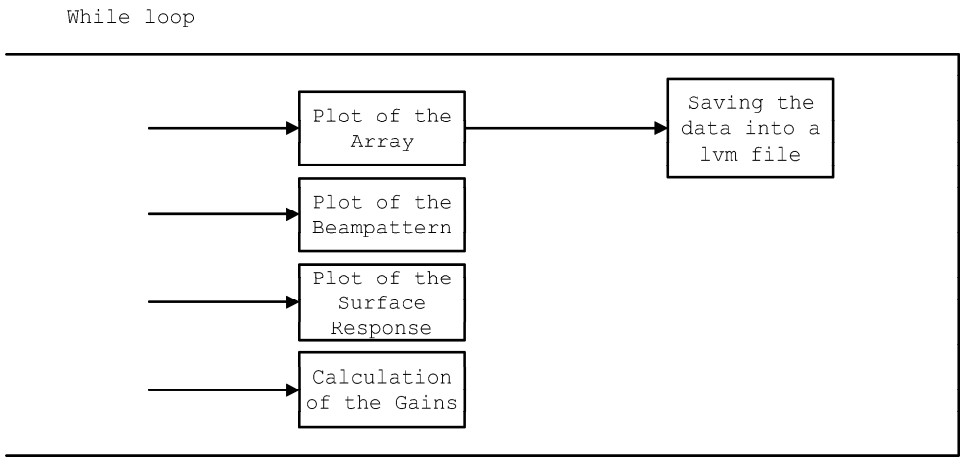

Figure 12: *The figure illustrates a block diagram of the final stage of the simulation*

2.5 The front panel, handling

As can be seen in the figure 13 the simulation front panel is seperated into array configuration and the simulation results section. In the array configuration section the array can be defined on its position and rotation around the axis. With the position there are 2 different posibilities, whereas the first on is on the common way → with array design, aperture size and amount of microphones. The other way is the definition by choosing the "'adjust coordinates"' button, with this way its possible to fill in the coordinates by hand or just variing it. With the calculation of the beampattern its possible to adjust the steering angle and to choose the desired frequency, furthermore spatial windowing can be applied by changing the weighting of the microphone channels. As mentioned from the principle its of course possible to plot the amplitude of the beampattern and the phase → this is done by pressing the „Phase/Magnitude“Button.

2.6 Improvement and other realisations

After working with beamforming systems the similiarty with FIR Filters became obvious. It has been tried to develope such simulations by just transforming the weighting values of the channels, what means an array with ones into the directional domain. The directional domain can be calculated by the 2D FFT and results in a Sinc function. With a coordinate-transformation the sinc-function can be pulled on a spherical surface and results thus in a beampattern. But the problem is the definition of the distance between the microphones or the spatial period because it is only possible to define arrays with regular distance. This principle is maybe a way to accerlate the software execution (if its still usefull) by using standard 2D FFT (or spatial convolution).

Another point, which is maybe interesting is the realization of a broadband version with frequency dependent weighting values to develope systems with constant beamwidth or to evaluate

the „gain quantities“with broadband beamformer. Further ideas due to the improvement will be just listed in keywords:

- Observation of the spatial aliasing limits with irregular arrays,
- Including different microphone characteristics (cardiod, super-cardiod etc.),
- Possibility to define boundary layers,
- Possibility to define different sources (not just point sources) etc.

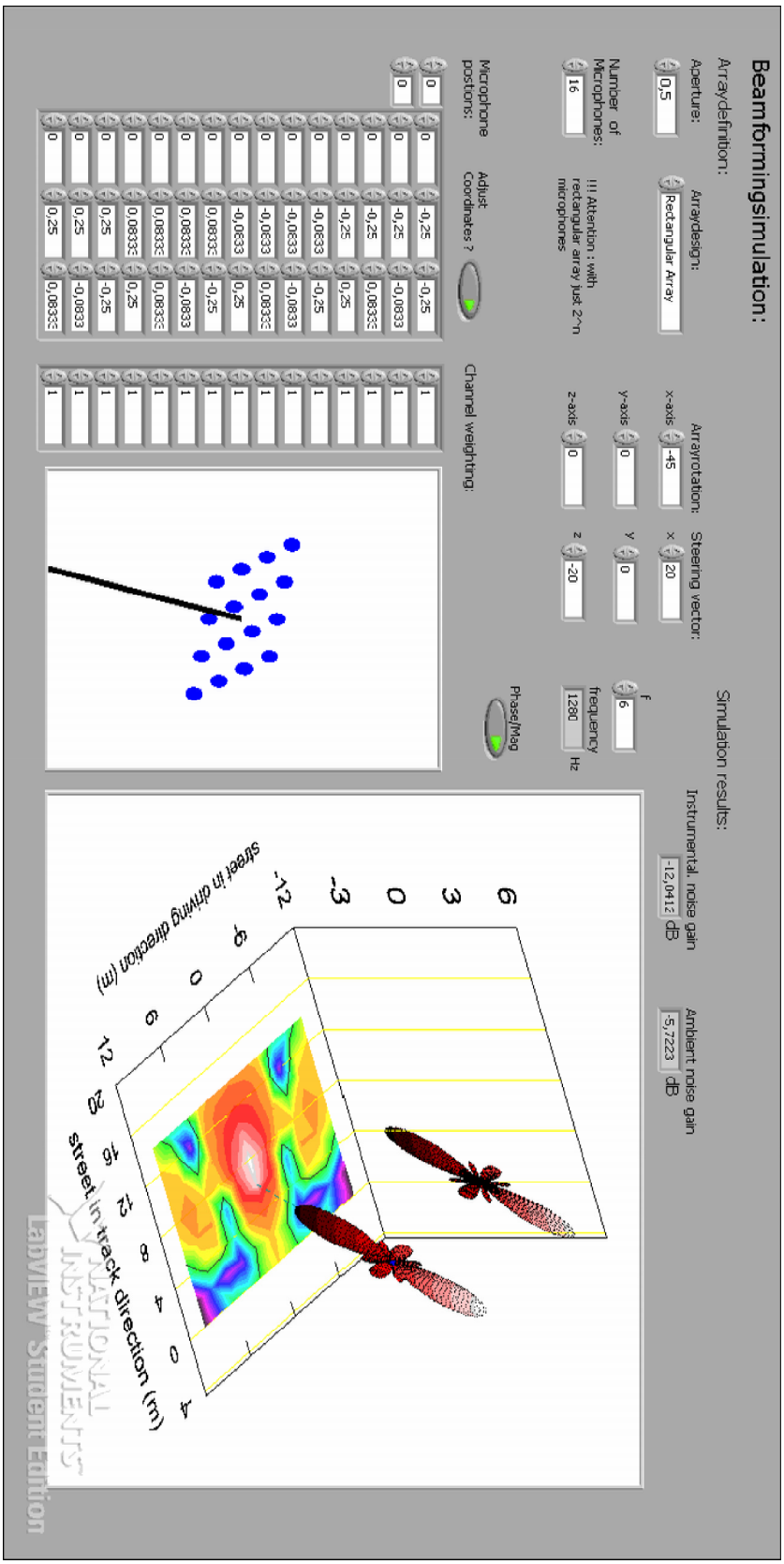

Figure 13: *The figure shows the front panel of the simulation tool in Labview*

Chapter 3

Analysis of the beamformer and the acoustic situation

Because the spatial filter is the most important part of the detection system, there has been a more precise look on the beamformer performance and additionally on the sound source. The sound sources are of course the vehicles (car, transporters and anything else) within their surrounding environment, the highway.

The analysis of the sources has been done in the first part of this chapter with the focus on finding out the noise behavior of the vehicles and the effects occurring with the street area. The second part will handle with the beamformer performance analysis and the impact due to the results. In order to show the effects, whose occurring with the measurement results the developed simulation tool is used.

3.1 Analysis of the acoustic situation

In this section the car within the environment, the street is analyzed in more detail. Cause the knowledge of the source behavior is essential to understand the beamformer performance in the end of the chapter.

3.1.1 The car as a sound source

Due to the fact that the car is localized by its noise characteristic, there should be a precise look on the car as a source within its area, the highway.

> *A car is a cluster of vibrations- and noise problems with a few parts of iron sheet on it, by Pilgrim P.* [*1*]

As the citation reveals is a car composed by a couple of sound sources, whereas the noise and vibrations are transmitted by the car structure, radiated to the car interior and outside into the ambiance.

In this work just the out sided noise radiation is of interest due to the fact that this noise is used to localize the vehicle with the system. This outside noise can be separated into two

main parts, listed below, whereas both parts having there own special frequency characteristic and dependent on the speed different power. Thus after [1] a vehicle with low speed radiates most of its noise through the intake/exhaust aperture. On the other hand with high speed the rolling noise is dominant. Below the two main parts, with their particular components are listed [1]:

1. The drive system noise:
    - drive unit noise (motor, gear drive, ancillary unit)
    - the intake/exhaust system noise
2. The rolling noise
    - tire and pavement contact noise
    - aerodynamic noise of the car body

In order to demonstrate the radiation behavior and the frequency characteristic of the vehicle noise, signal analysis (time, frequency and spatial) is done with two passing-by situations:

1. A **diesel transporter with 10 km/h** : The transporter passes through the measurement array mounted on a mobile gantry. The noise is analysed by a spectrogram (additional a periodogram via welch at the crossing points) and hearing analysis (see figure 14). With this techniques it turns out that the noise is caused, as mentioned at low speed, by the motor system with its power in the low frequency area. Furthermore the exhaust and the intake aperture are the most dominant sources of the motor system and beside this coherent, what is important due to loudness and interference.

   With the spatial analysis (see figure 15), or in more detail the directional time analysis, a precise look on the radiation characteristic of the vehicle is done. As with most complex sound sources (for example music instruments) is the radiation directional and frequency dependent. By considering the result of the direction time analysis it can be seen that the line, which represents with its incline the driving direction, is broken from the middle. This „break of the line“is due to the fact that the transporter, after crossing the middle (or the 0° beam of the array in driving direction) is not acoustic visible anymore → the motor radiates less energy in the back (the exhaust aperture noise is too low). Because of this information leak and the fact that even with high speed the vehicle is acoustic visible just from ahead the solution to mount the array above the vehicles is not recommendable.

2. A **passenger car with 140 km/h**: In the second case a passenger car is passing by with 140 km/h across the sidewards (on the gantry) mounted microphone array (figure 14 and 1), thus with much more speed than in the first case. Hence the spectrogram shows a more broadband signal (which is due to the fact that the noise level is strongly dependent on the speed). From the literature it is known that with high speed the „tire pavement contact noise“is dominant and beats the motor system noise. Again with the direction-time analysis it can be seen by looking sidewards on the car even noise from the back is present what underlays the assumption that sidewards the car the tire pavement noise is dominant.

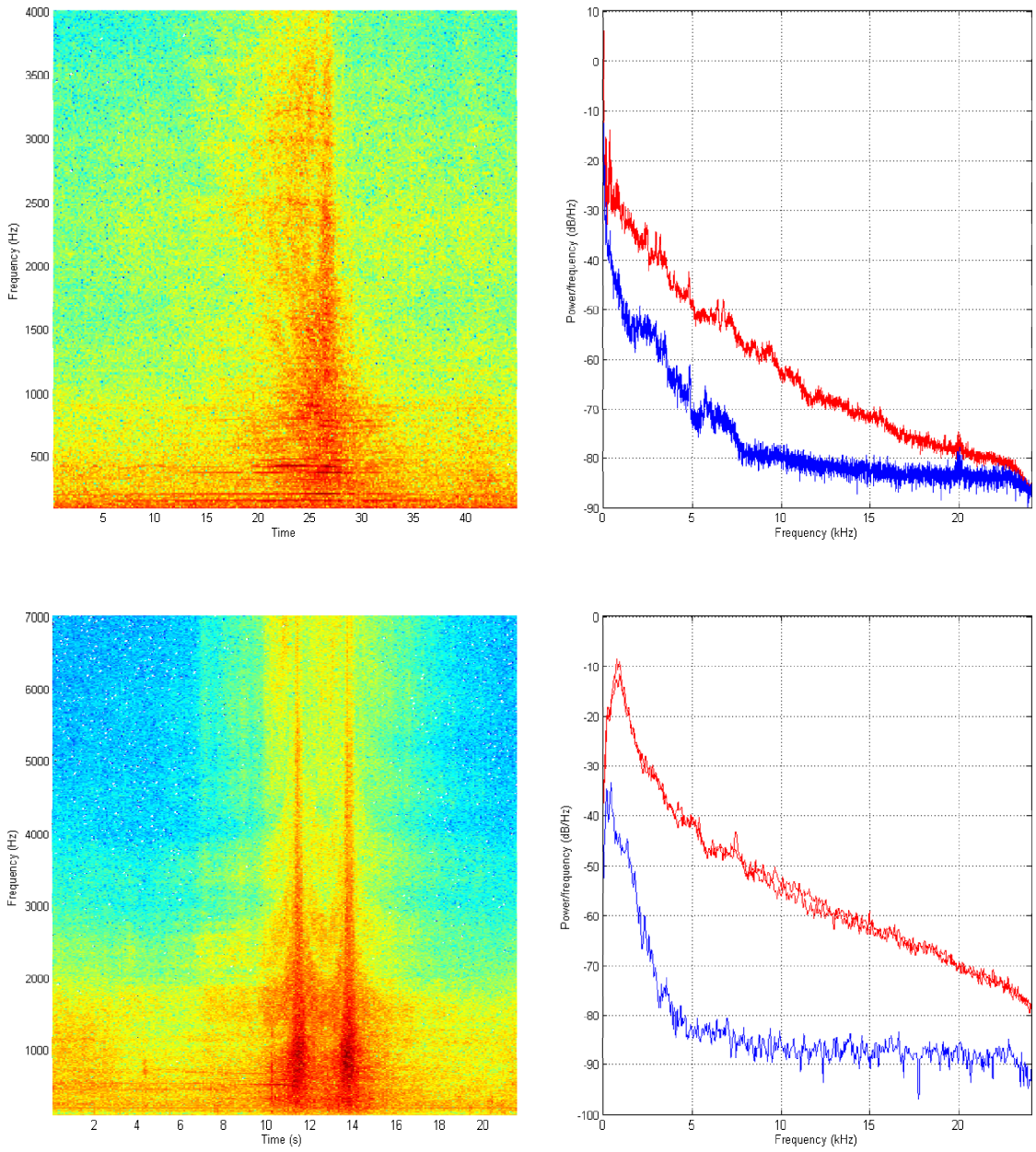

Figure 14: *The first spectrogram clearly shows a change in the spectral energy during the crossing by the transporter. The right plot demonstrates the psd estimation (via welchs method) in a silence situation (blue) and during the crossing (red) → the difference can be seen as a kind of SNR. The meanings of the plots on the bottom are similar to the first ones but with 2 car crossings*

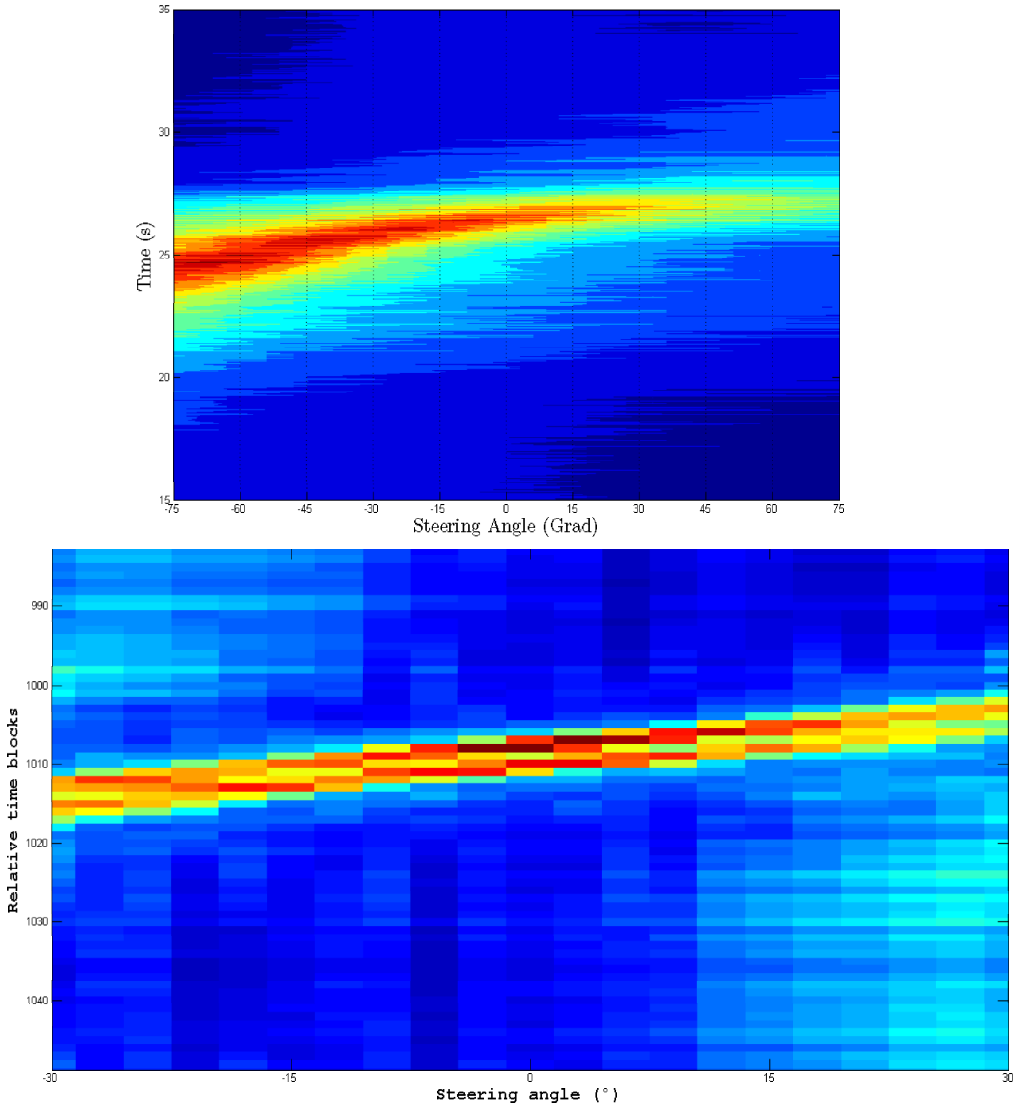

Figure 15: *The directional time analysis, once for the low speed transporter (and the above mounted array) and once for the high speed car (with sidewards mounted box)*

Finally it turned out that the noise is in fact a composition of the two main sources, whereas with for example high transporter carbins the aerodynamic component arises as well. This fact is confirmed by the figure 16 from [1] which shows a crossing car with 140 km/h.

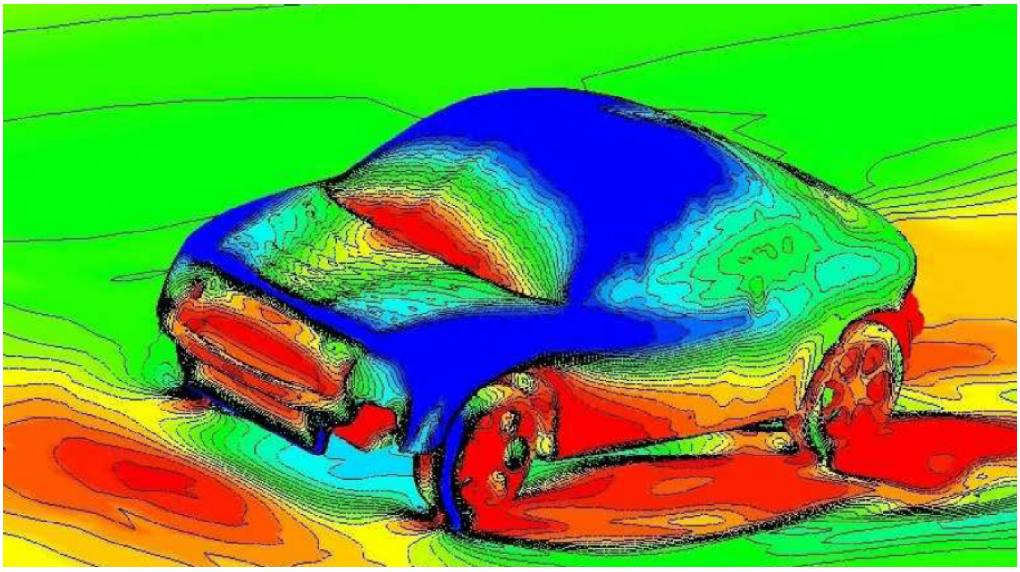

Figure 16: *Radiation characteristic by a passing by car with 140km/h [1]*

3.1.2 The acoustic road situation

Due to the challenges concerning the acoustic street situation this section is divided into three areas:

1. Reflections and Reverberation
2. Interference
3. Spatial noise

Reflections and reverberation

In this part the street and the surrounding with the impact on the measurements will be discussed. By considering the street as an acoustic area it turns out that there is just one reflecting surface → the street surface. In fact there are of course a few reflections (from the gantry, traffics sign, other cars etc.) but the more relevant ones are the first reflections. In the case of a close noise barrier or other boundaries it should be talked about reverberation.

In the „Room Acoustics“theory, reflections are described by the mirror source principle. Thus with a single reflection on the street surface the car can be seen mirrored on the street surface. The strength of the mirror-car is of course dependent on the absorption of the street surface → for example with whispering asphalt the absorption is really high.

The problem with beamforming due to reflections is that the real source and the "'mirror"' source can not be distinguished. That means with a real reverberant room beamforming becomes a real challenge.

In fact with our system, there is just a first reflection on the street surface and thus a little drift and broadening due to the real source positions (figure 16). Therr are a few methods in literature to eliminate mirror sources (coherence filtering etc.).

Interference

Beside the sound source there are the competing sound sources or just other vehicles, whose are called interferences. For that reason in the case of more than one car in the observed area the problem is not anymore just a single source location problem → multiple sources occurring. The task to localise more than one sources would not be a problem or a challenge if the spatial filter would be ideal. But in reality there is the beam pattern with its side lobes structure and leads thus to the so called „ghost images"→ it can't be distinguished between a real source or just a side lobe. For better understanding a numerical example is shown:

Acoustic situation:

1^{st}lane: a transporter with 75 dB
2^{nd} lane: a passenger car with 65 dB

The used array is a line array with 8 microphones and thus the maximum side lobe level (which is only dependent on the number of microphones results in:

$$A_{sidelobe} = \frac{1}{M\ \sin(\frac{3\pi}{2\,M})} = \frac{1}{8\ \sin(\frac{3\pi}{2\,8})} = 0.225 \rightarrow -12.3\,dB \tag{3.1}$$

As mentioned above the problem is the separation between the side lobe and the 2^{nd} source. This separation is just possible with a greater second source level than the side lobe level of the first source (the transporter). In this example the passenger car delivers 65 dB and the side lobe level of the transporter results in 62.7 dB. Therefore the car can be localized theoretical because with just 2.3 dB it will be a challenge.

In general, the main side lobe level difference is also called „plot dynamic"and indicates therefore the possibility to separate sources with a specific level difference → dynamic. Furthermore since the occurring of side lobes with beamforming can be seen as a spatial leakage spatial windowing can be applied as well.

Spatial noise

There are three types of noise with multichannel system, whose are different in the degree of correlation between the channels. The spatial correlation is measured by the spatio coherence function (see [4]):

$$\Gamma_{ij}(f) = \frac{\Phi_{ij}(f)}{\sqrt{\Phi_{ii}(f)\Phi_{jj}(f)}}$$

$\Phi_{ii}(f)$ is the cross spectral density between signal i and j.
The coherence is essentially a normalised cross-spectral measure, as the magnitude squared coherence can be seen to be bounded by $0 \leq \Gamma_{ij} \leq 1$.

Incoherent noise fields: With $|\Gamma(f) \approx 0|^2$ the noise measured at any given spatial location is uncorrelated with the noise measured at all other locations. Such an ideal incoherent noise field is difficult to achieve, however in the case of a microphone array the electrical noise in the channels (preamp, quantisation noise etc.) can be considered to be a source of incoherent noise $\rightarrow$ spatial white noise field.

Coherent noise fields: $|\Gamma(f) \approx 1|^2$ indicates the coherent noise field, whereas the noise measured at any given spatial location is uncorrelated with the noise measured at all other locations. In practice, coherent noise fields occur in open air environments where there are no major reflections.

Diffuse noise fields: The diffuse noise fields are in-between the mentioned fields. Many noise environments can be characterised by a diffuse noise field, such as offices, car noise etc. The coherence therefore can be modeled by equation 3.2 and is dependent on the microphone distance and the frequency.

$$\Gamma(f) = sinc\left(\frac{2\pi f d_{ij}}{c}\right) \tag{3.2}$$

With this coherence function its possible to rate the performance of the standard sum and delay algorithm because the sum and delay algorithm works best with completely incoherent noise (the plot in figure 17 shows the coherence function with the smallest distance of our array dependent on the frequency). By considering the plot it becomes obvious that with low frequency the coherence functions becomes one and thus the ordinary sum and delay beamformer delivers no SNR improvement compared to a single microphone [3].

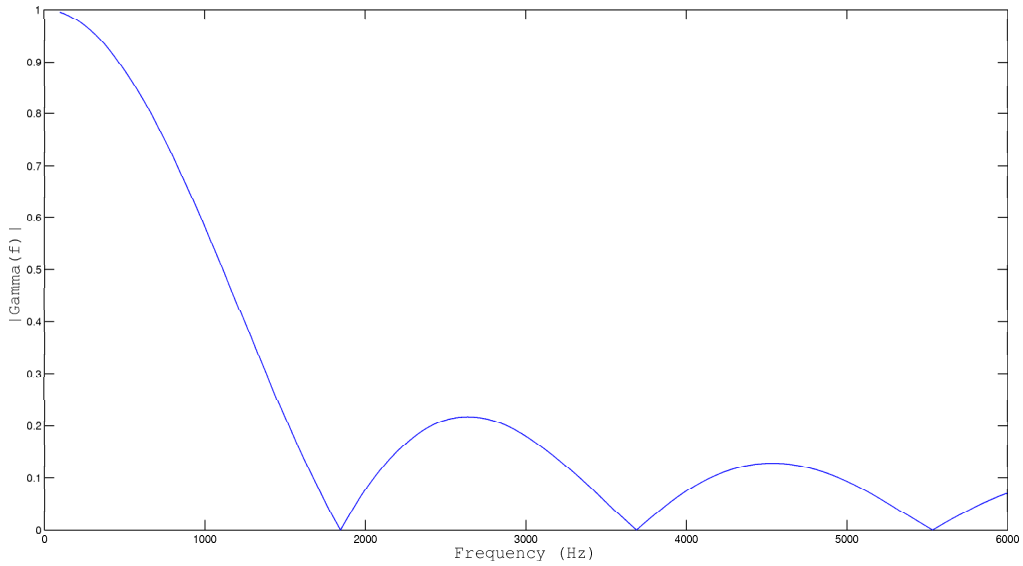

Figure 17: *The figure shows the coherence function on frequency*

3.2 Analysis of the beamformer

In this part their will be a more precise look on the beamformer performance and the effects occurring on the measurement results. Furthermore with the results the requirements to the system will be mentioned.

3.2.1 General requirements to the beamformer

In general within the beamformersystem it should be possible to separate the noise sources in two ways:

1. Separation of vehicles in a row
2. Separation of vehicles side by side

Bot requirements refer to the main lobe width, one in driving direction and one in track direction. Additionally the maximum side lobe level should be smaller than the dynamic between the vehicles (see figure 18). From images processing it is well known that an object can be resolved just by the size of the point spread function (PSD) which is similar to the beam pattern or the main lobe width in acoustics.

The analysis itself is done by taking the directional-time responses during a car is crossing the array. Both line arrays have been considered apart:

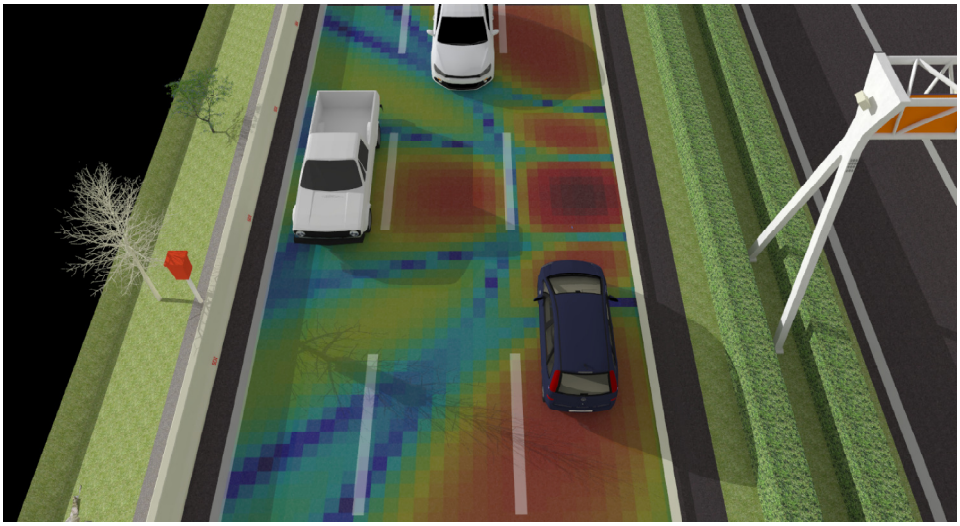

Figure 18: *The figure shows the situation on the highway with an steering example thus one main lobe used to separate the vehicles*

3.2.2 X-array

The x-array is the linear array scanning up the street in driving direction (the array is positioned parallel to the street surface). For the calculation of the time-direction analysis the array is steered to the desired directions. But of course this steering is not done after another but rather at the same time. Followed by this calculations the result is temporal filtered an the energy calculated. This result represents the base for further calculations and considerations. For understanding the results it is of course essential to know the direction and the steering definition across the street. This is clarified in the video snapshot (figure 19), done synchron with the audio data aquistion, by the arrows.

The next figure 20 illustrates the above mentioned response of a beamformer in two different ways. Once there is the plot with a few lines one it representing the spatial result for a few cars over a large duration. In fact the lines having different incline due to their speed and direction. For the second plot in the figure zooming is used to demonstrate „a vehicle crossing“on a more particular way.

Vehicle crossing

To understand the behavior or the performance of the x-array the result in figure 20 is used. Therefore the vehicle crossing will be explained step by step.

After the vehicle is reaching the steering area the energy (which can be seen as color) will increase. But of course the energy increase is depending on the beam pattern for the particular direction and the sound source itself. Whereas with large steering angles the car is captured mainly from the front/back and thus the motor noise/exhaust noise is more present in the result. By considering the further movement of the vehicle it will cross every particular steered beam pattern and thus the line (seen in the figure 20) will occur. That means if a vehicle is

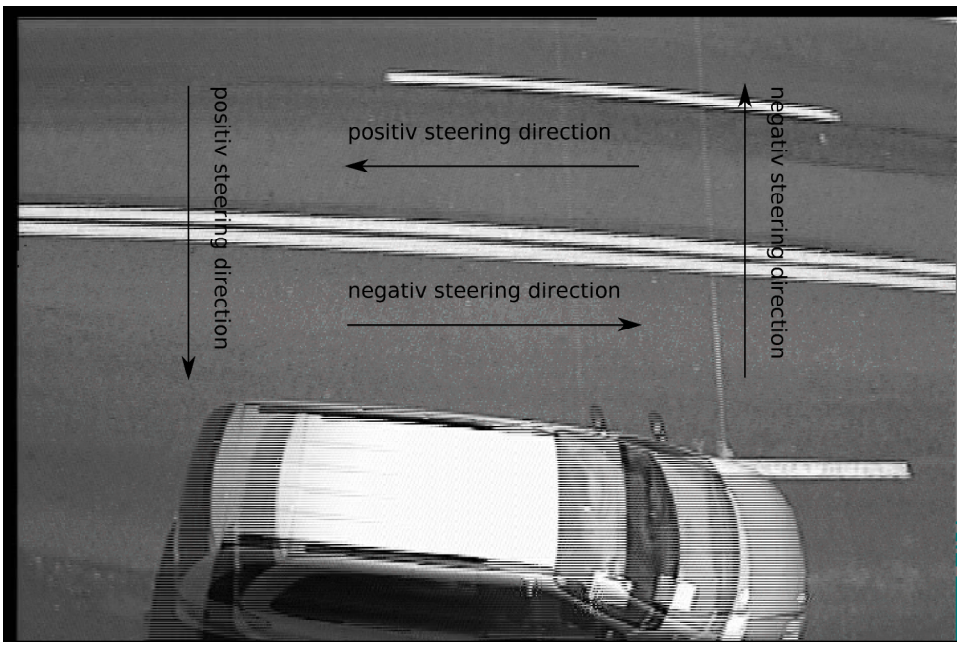

Figure 19: *The definition of the steering / driving directions*

crossing the x-array in a duration of time with special steering angles the direction and the speed is completely known.

One would expect the line is equally wide but the width is not the same for every particular direction. This fact can be explained by having a look on the theory part again, by considering the behavior of the main lobe width with steering. Because as derived in chapter 1.4.5 the main lobe width off axis is dependent on:

1. The frequency
2. The steering angle

That means the change of the line width can be explained by the main lobe width dependency on the steering angle, which has been $1/\cos(\phi)^3$. But of course the second parameter, the source (a combination of the tire and the motor noise) will effect the main lobe width as well.

By sighting the figure 20 again, another effect can be seen in the outer steering areas, once below the mentioned line and once under it. The areas are highlighted cause there is more energy than in the surrounding area. Since there is the aliasing effect in space as well it should be excluded by an aliasing check. Additionally the plot 21 demonstrates obvious the decrease of the nyquist frequency with increasing steering angle.

Simulation results

The simulation is used to prove the assumptions worked out in the part before. Therefore just the street response results are taken for special steering angles.

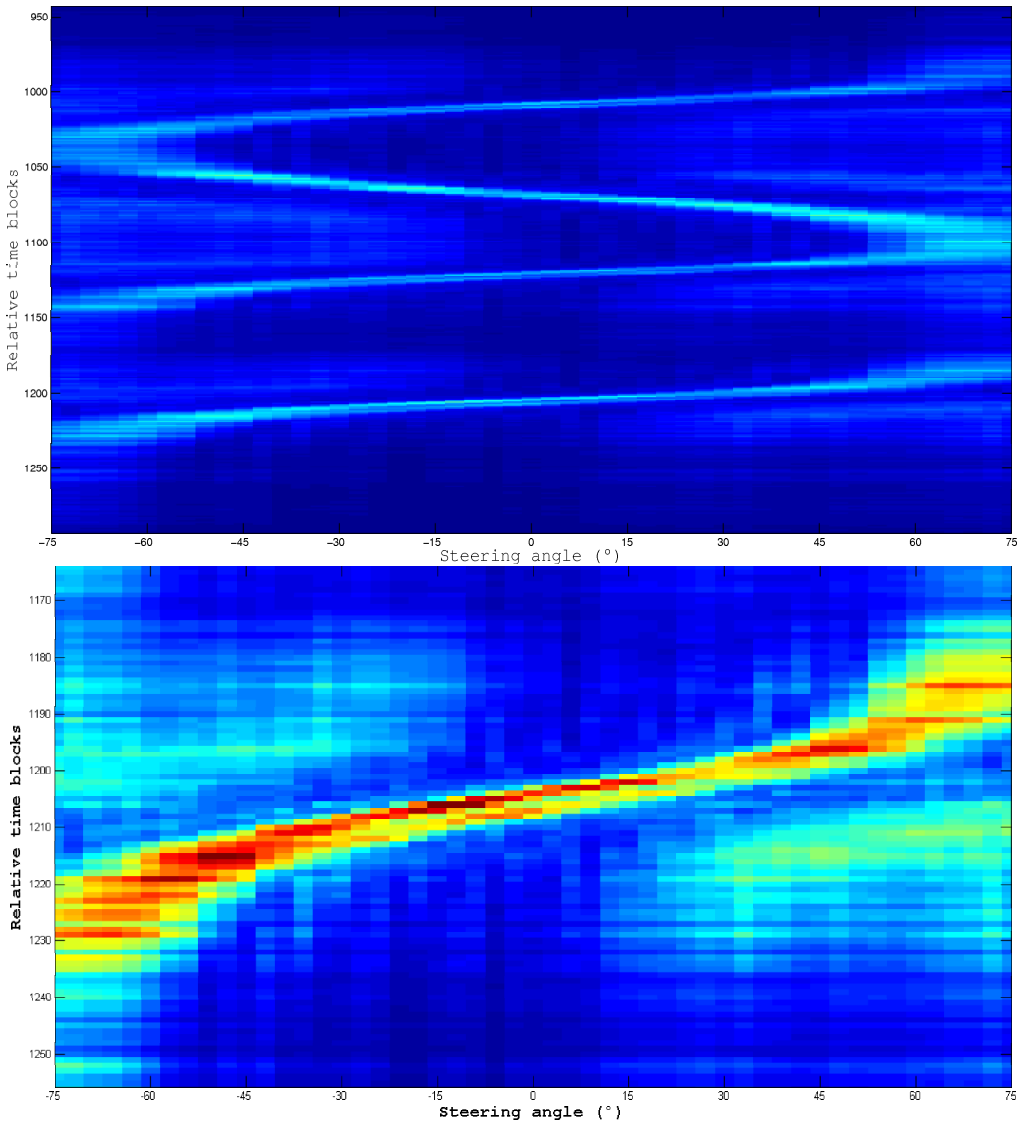

Figure 20: *Radiation characteristic by a passing by car with 140km/h*

The beam pattern response for the street surface with the 0° steering can be seen in figure 22. By regarding the beam width it is getting apart that there is no zero crossing, which defines the main lobe width and thus the spatial resolution the filter. But this problem is solved just by taking the minima between the main lobe and the first side lobe, thus an approximation is done. In this case it turns out that the main lobe width is increasing with the distance (y-axis) as well not only with the steering angle. The greatest side lobe level, which defines the dynamical resolution is ≈ -12 dB and constant through the distance on the y-axis.

The next figure illustrates the beam pattern for 30° steering 23 and will be considered as well. In this case the main and side lobes are shifted and distorted due to the distance on the y-axis and the steering angle. The level of the side lobes is constant as well with −12 dB but with the outer area the beam response is increasing again. This effect is due to spatial aliasing (no side lobe) and therefore the prove for the assumptions before is done.

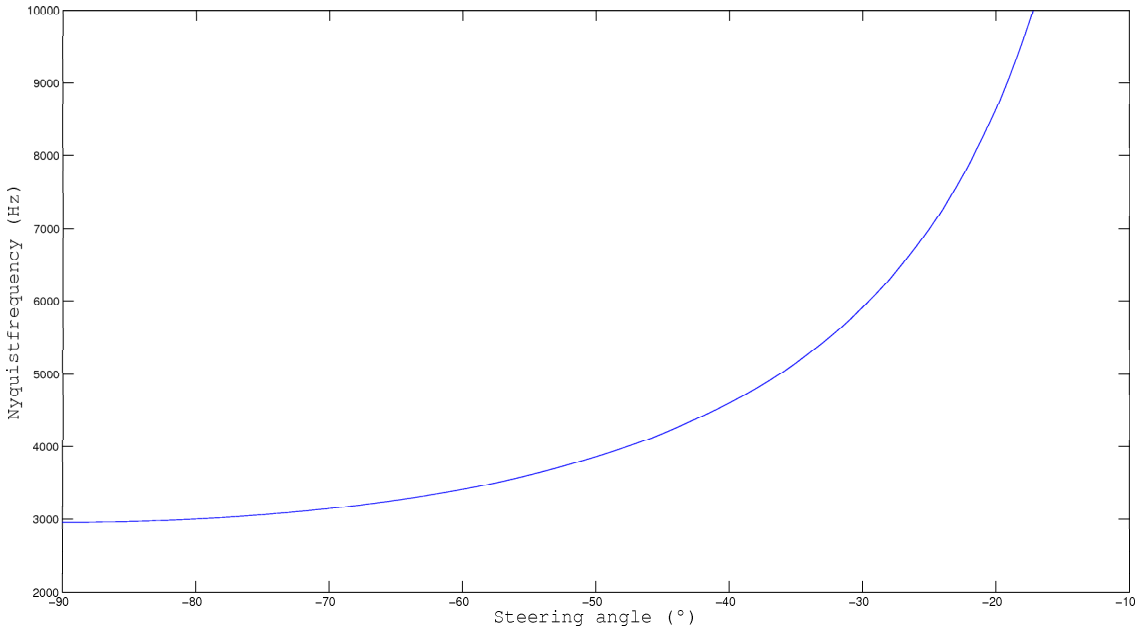

Figure 21: *Nyquist frequency dependent on the steering angle*

Finally some values for the beam width (0° and 30° steering) are measured, wherase the values are approximations:

distance on y-axis (m)	**resolution 0°**	**resolution 30°**
0	1 m	2.5 m
5	1.3 m	3.3 m
10	2.5 m	4 m
15	3 m	4.5 m

Even if the resolution is with 30° (see section 1.4.5) not double as large than with 0° (cause its approximately measured) the dependency by the off axis resolution is proved.

Energy resolution

By considering the spatial resolution, the afterwards temporal processing has to be considered as well. Because as mentioned the spatial filtered signal is afterwards processed and the energy calculated. Thus this results in blocks of energy values (for every particular direction) quantized by the window width. But what is if 2 cars are really close after another like in the figure 24 and a window is two times as large as both windows (green and the red window) just one source can be localized. On the other hand if the window is to small the energy will be spread through a few windows and thus the SNR becomes worse.

Thus defining the optimal window size is a challenge cause the duration of the sequence which includes the most car energy (spatial filtered) is dependent on the speed and of course dependent on the beam width. For demonstration two examples are calculated:

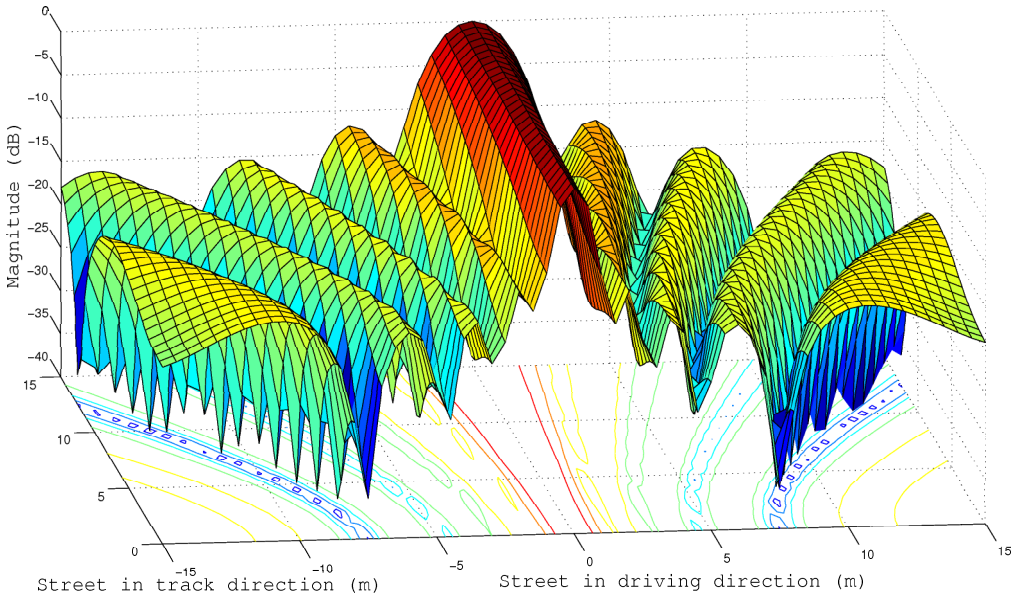

Figure 22: *The x-array simulation result for 0°*

1. a car with 150 km/h on the 1^{st} lane and thus with 2m beam width results in a window of 72 ms
2. a car with 90 km/h on the 3^{st} lane and thus with 6m beam width results in a window of 240 ms

Figure 25 shows how the different window sizes effects the signal. It results that with the right window size even two peaks or sources can be recognized whereas the SNR is only best for the right fitting window. The consequence of overlapping can be seen in figure 26 as well whereas the line, representing a car is more narrow compared to the line in figure 20 and thus easier to use for detection.

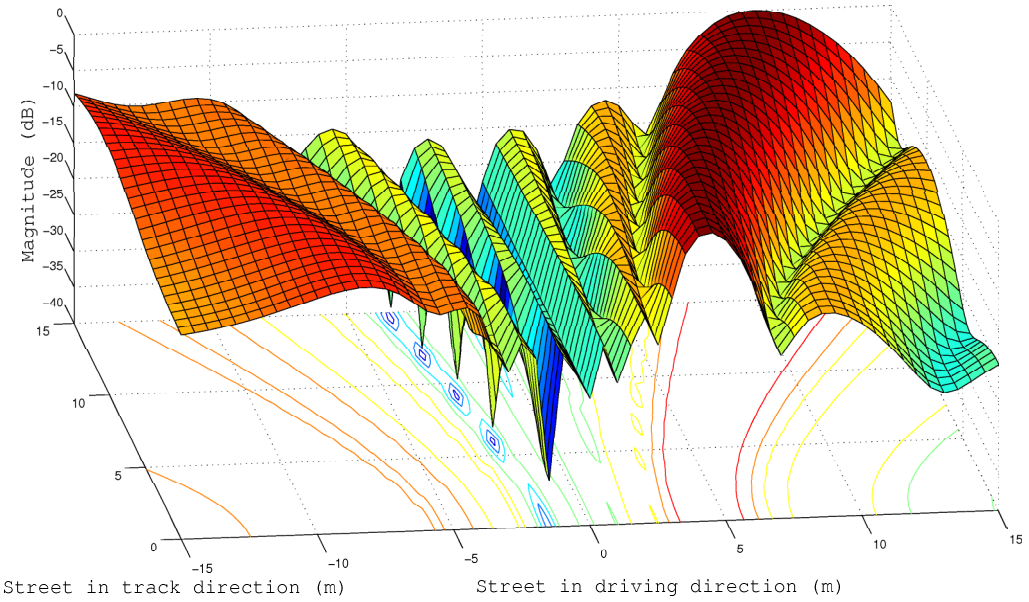

Figure 23: *The x-array simlation result for 30°*

3.2.3 Y-array

The y-array is also a linear array perpendicular positioned on the x-array. Additionally the whole box including the array is rotated with -45 ° against the gantry (see figure 1). Thus the 0° steering should match with the center of the lane. Again the result of the y-array measurement is taken to analyse the behavior of the system.

Vehicle crossing

As like before the vehicle crossing will be explained in more detail. In this case it is useful to choose a few examples on different lanes to get feeling for the effect occurring. The measurement process with the y-array it is almost the same as with the x-array thus the array is steering up the driving surface in track direction. The example results can be seen in figure 28 whereas their are 2 cars on the 1^{st} lane (next to the gantry) and one on the 4^{th} one in opposite direction. The center of steering is again on the center of the driving lane between the 4 tracks.

It could be expected that there are just highlighted areas representing the cars on the lanes but there are distortions, especially on the 1^{st} track. Additionally to the distortions due to the beam pattern there is the distortion of the projection surface. This fact can be seen in the figure 28 by the white lines separating the tracks.

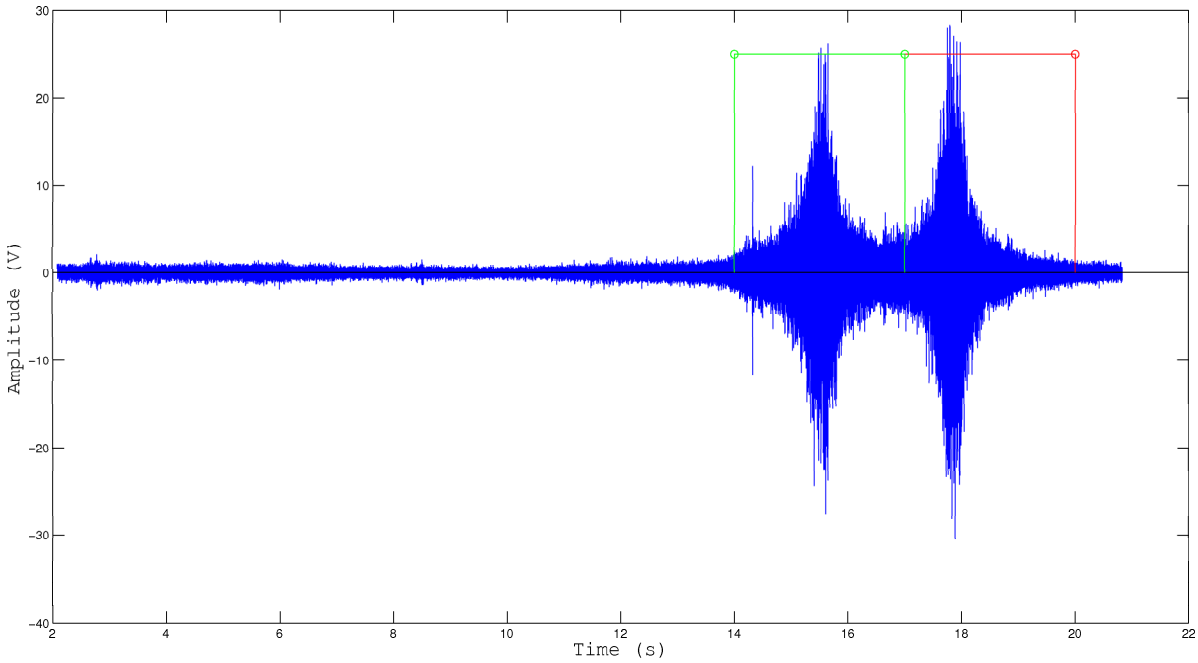

Figure 24: *Signal sample to demonstrate the energy quantisation*

Simulation results

To understand the curvation, especially with the cars on the 1^{st} lane the simulation tool is used again. The simulation is done once with 0° (means in the direction of the surface center → the double solid line) and once with 30° in gantry direction. By the use of the simulation results it can be seen that the curvation is due to the steering distortions of the y-array. With increasing steering angle in gantry direction the main lobe curvation increases → in fact the car is recognized two times by crossing a „ring“main lobe. Finally from 30° the main lobe becomes really hard curved and thus is only sensitive in a small area. Another fact,which can be seen in the figure is the constant maximum side lobe level, even with steering it does not change.

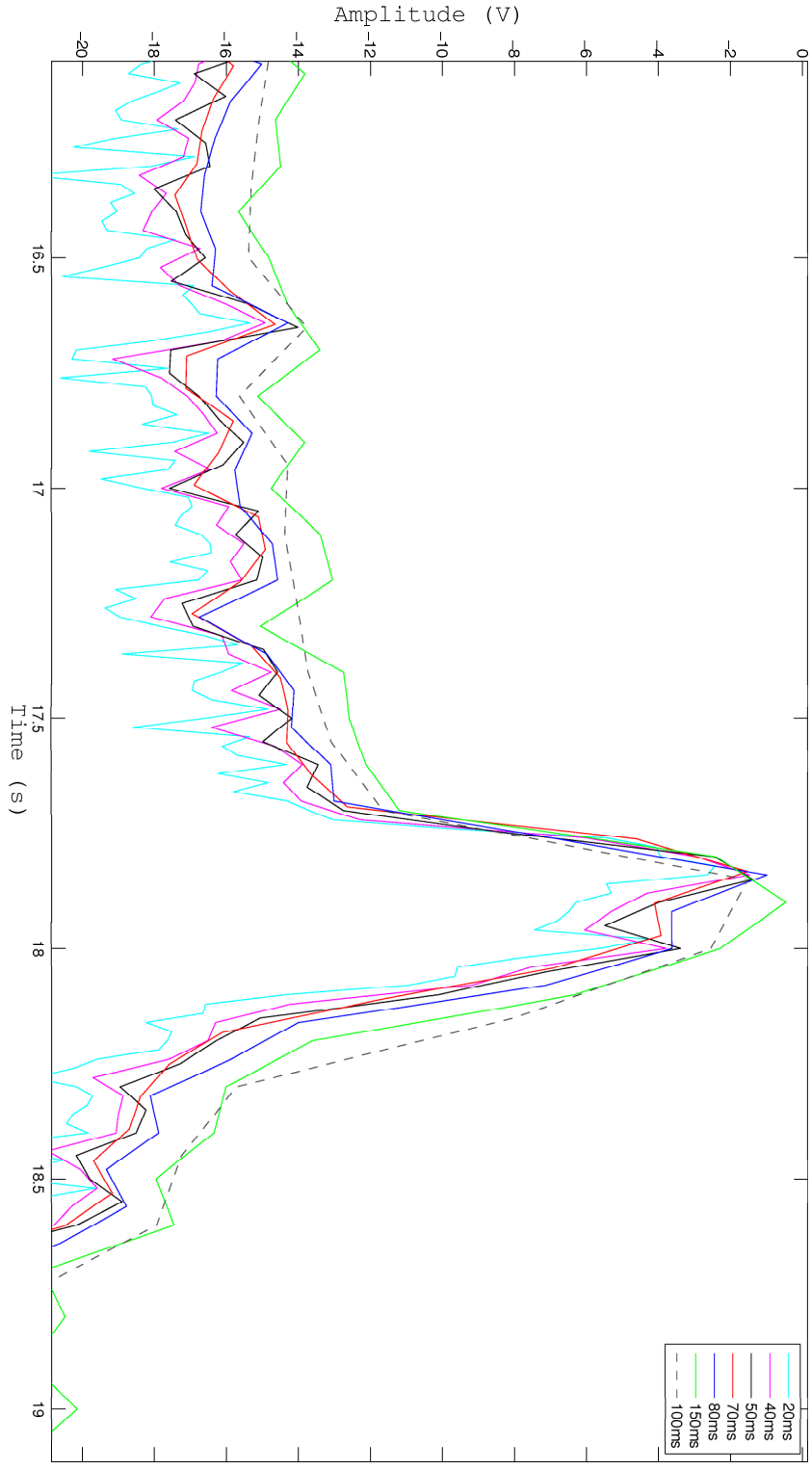

Figure 25: *Different windows and their effect on the signal*

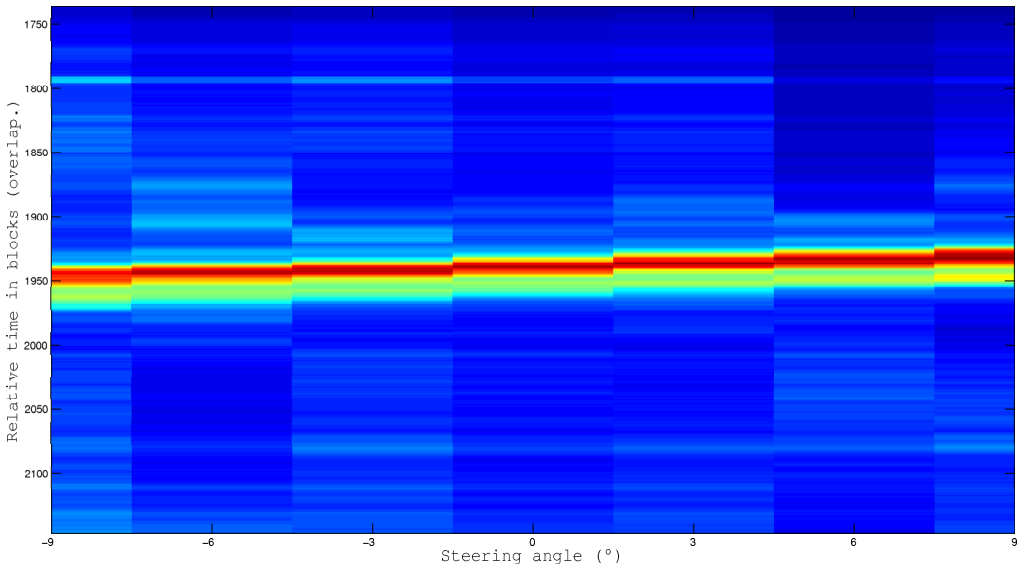

Figure 26: *The effect of windwos overlapping on the direction line*

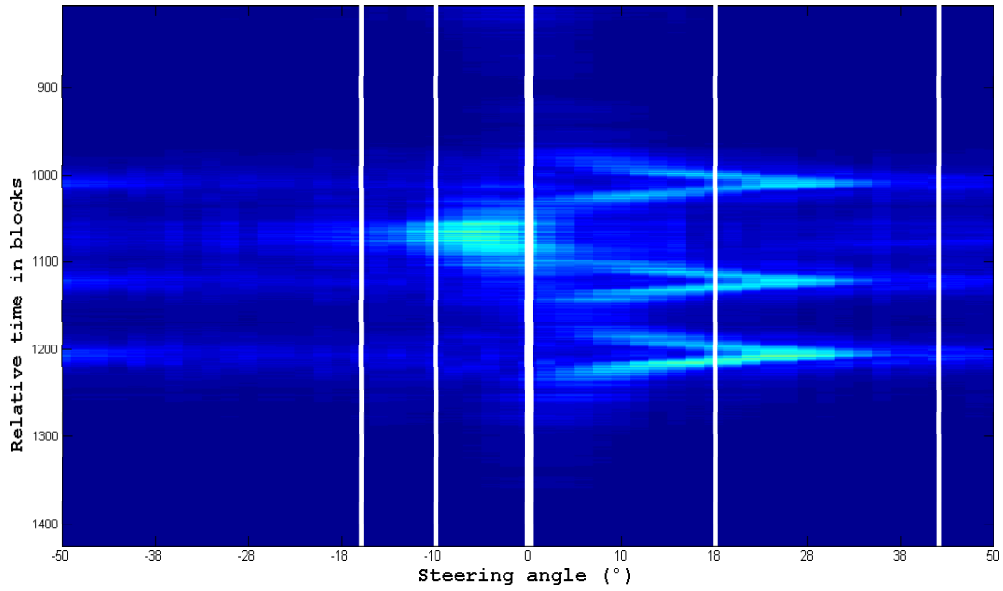

Figure 27: *The measurement results by the y-array*

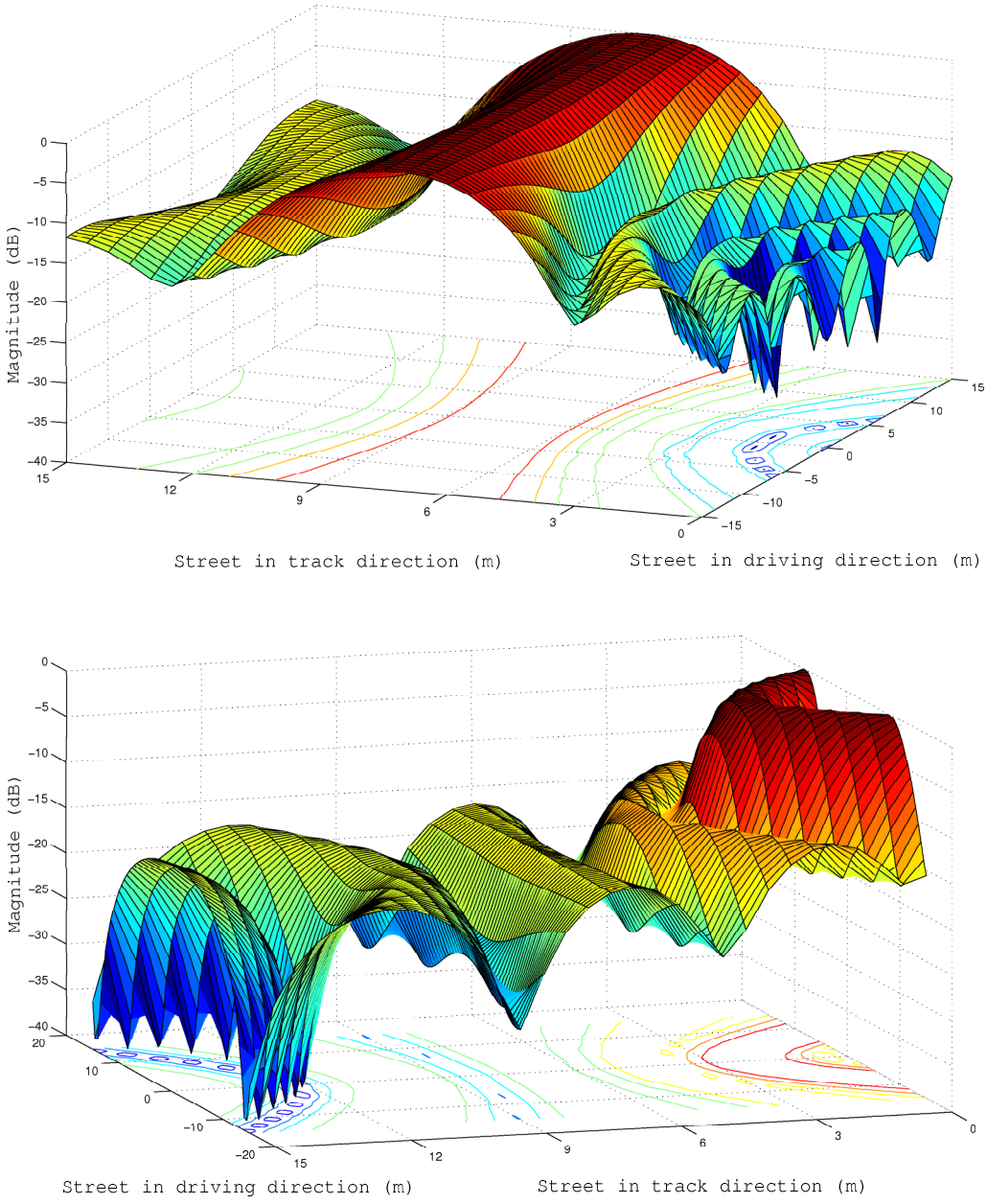

Figure 28: *The y-array simulation response for 0° and 30° steering angle*

Chapter 4

2D Array development

4.1 Array design

4.1.1 Irregular sampling

As mentioned before in this chapter it will be handled by the development of a 2D array. At this point it should noted again that in space its common to sample irregular. The derivation or the specification of the optimal sampling positions is done by defining the „array design“.

To show the advantages of irregular sampling a line array is used, whereas once the distances between the microphones are equal and once non-uniform or irregular.

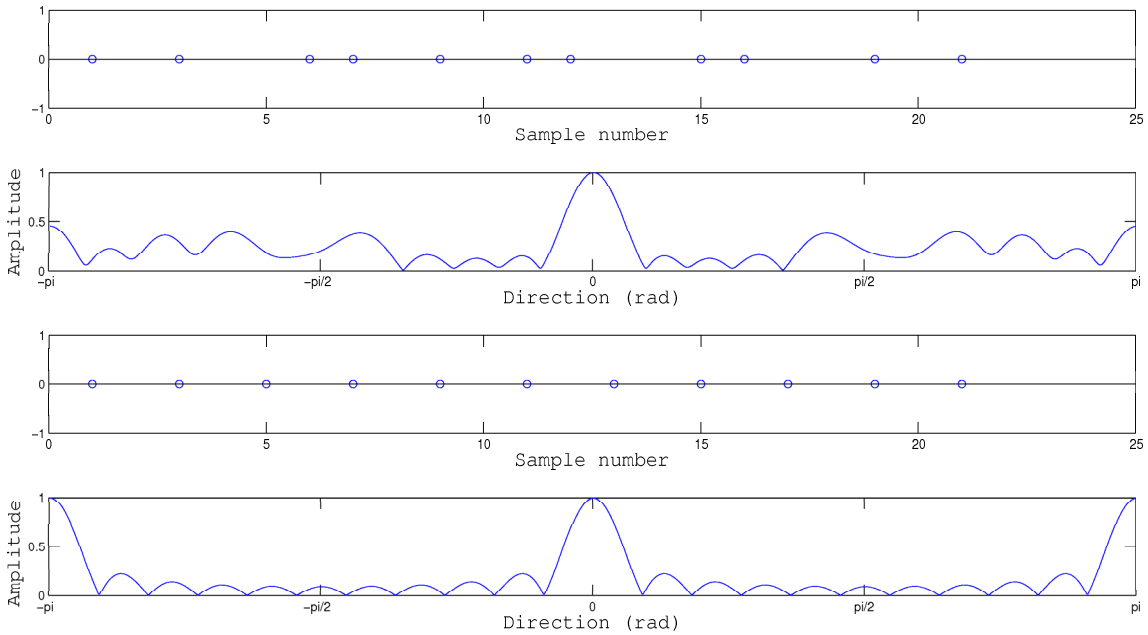

Figure 29: *Effects on uniform and non-uniform sampling*

By considering the results of the simulation, especially with the uniform sampling the periodic continuation of the spectra can be seen at $\pm\pi$. In the case of aliasing this peaks will move into

the range of usage and therefore will distort the result. In the case of non-uniform sampling the periodic peaks are more spread over the spectra and thus the aliasing effect will not occur in such an extend. But of course no advantage without disadvantage → the spatial SNR or the side lobe structure in sum is increased [5].

Additionally with irregular array designs the calculation of the quantities becomes more difficult, what leads again to the usage of the simulation tool.

4.2 The Cross array

The first array design trial was done by a cross array within the same coordinates as the x, and the y-array. The simulation results of this design can be seen in figure 30, whereas with two different ways of calculation (The letters are bold to clarifies the complex identy):

1. Arithmetical sum, the array is simulated in sum what relates to the sum of the complex values of the particular line arrays

$$\mathbf{B}(\omega,\phi)_{cross-array} = \mathbf{B}(\omega,\phi)_{x-array} + \mathbf{B}(\omega,\phi)_{y-array} \tag{4.1}$$

2. Geometrical sum, the arrays are summed geometrical instead (recommended by [2])

$$\mathbf{B}(\omega,\phi)_{cross-array} = \sqrt{\mathbf{B}(\omega,\phi)_{x-array}\,\mathbf{B}(\omega,\phi)_{y-array}} \tag{4.2}$$

By examining the results it turns out that the cross array shows, due to composition of the two line arrays, a cross shaped main lobe (is no side lobe structure). Both ways, geometrical and arithmetical sum, having their defect. The geometrical one shows a decrease in the main lobes but with increased side lobe structure. The behavior of the arithmetic calculated cross array pattern is vice versa. To compare both results the ambient noise gain can be used:

	Arithmetic	**Geometric**
G_a	-13.15 dB	-12.84 dB

Both results are of course just valid for a singular frequency (in this case ≈ 3 kHz was used) but in general the behavior can be estimated cause with increasing frequency the side lobe number will increase (to M+1, M is the amount of microphones) and therefore move into main lobe direction.

Finally it should be noted that within this design the source separation is even more problematic cause the dynamic difference between the maximum peak of the main lobe and the other levels is not enough. Thus the separation across the lane and in driving direction is hardly possible.

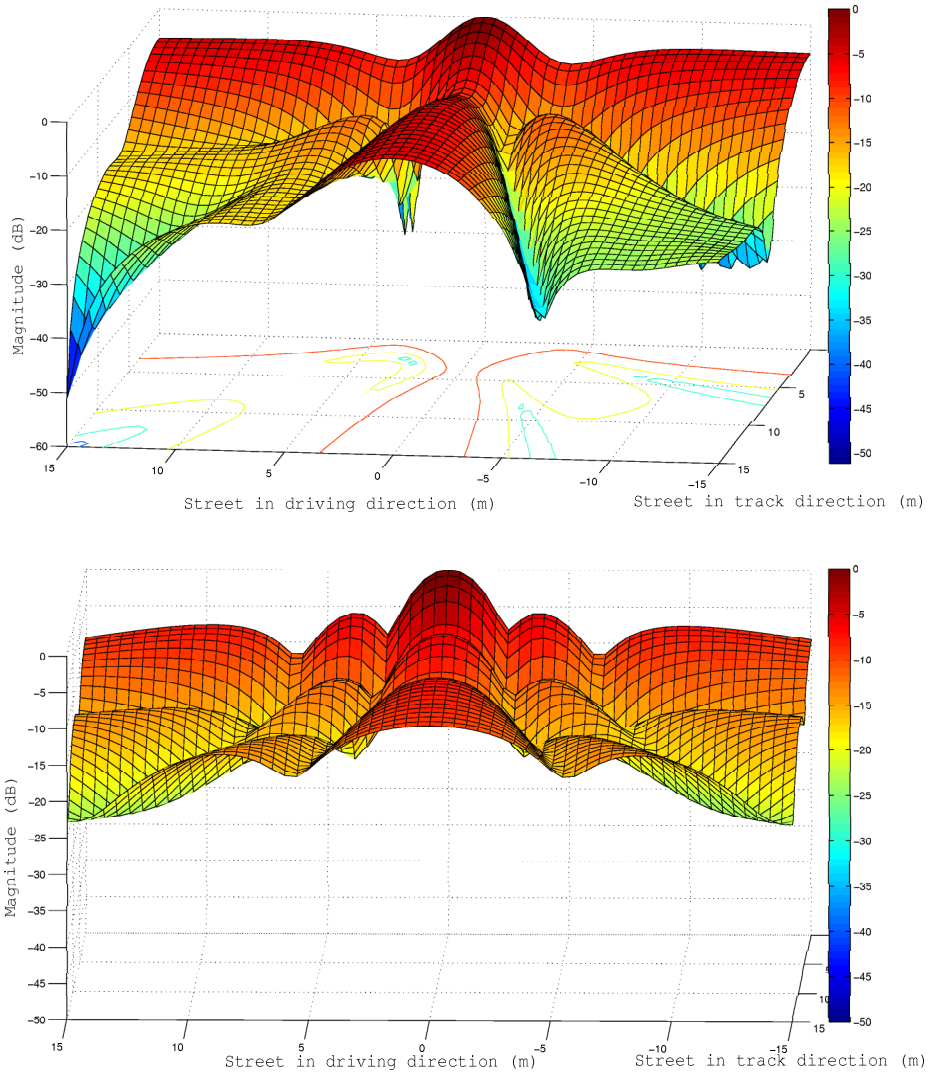

Figure 30: *The arithmetical and the geometrical calculated cross array*

4.3 Further array designs

In general it should be said that there are of course a few techniques to develop array designs for a specific situation. Like with temporal filters it is possible to design a spatial filter as well by its specifications. One method is for example the Parks-McClellan Filter Design [3], which is in fact a least squares method minimizing the error difference between a desired frequency response and the real one → this method can be used for spatial FIR filter as well but is of course complex and time consuming. Thus the more practical way is the trial and error method, by simulating a few designs and rating the results due to the desired situation.

4.3.1 The circular array

The circular array is by far the array design which delivers the best main lobe width. On the other hand the side lobes or the lobe structure beside the main lobe shows a circular behavior but with increased level. To rate the array design for the desired street surface the ambiance noise gain is -14.98 dB (3kHz).

With figure 31 it can be seen clearly that aliasing will occur, that cause the distance between the microphones on the circle:

$$2\pi r = 2\pi 0.3\, m = 1.88 \rightarrow 1.88/16 = 0.11m$$

and thus greater than the half wavelength for 5 kHz.

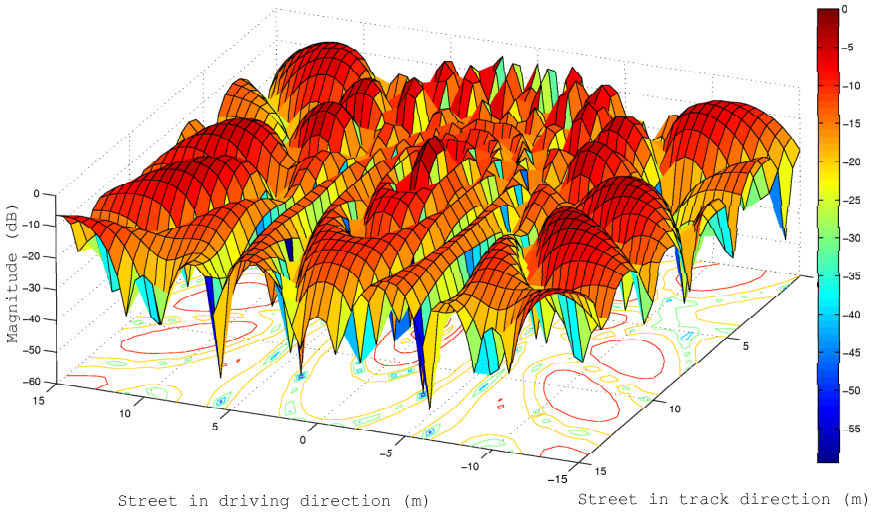

Figure 31: *The circular array design response*

4.3.2 The grid array

With the grid array its completely different, it shows a great structure around the main lobe but the aliasing components lying completely adverse in the lane direction. At this point it has to be stated:

With spatial sampling its hard to avoid aliasing completely thus aliasing should occur in directions whose not directly in the source field.

In our case that means the aliasing components should not move (with frequency) on the lanes or across the lanes.

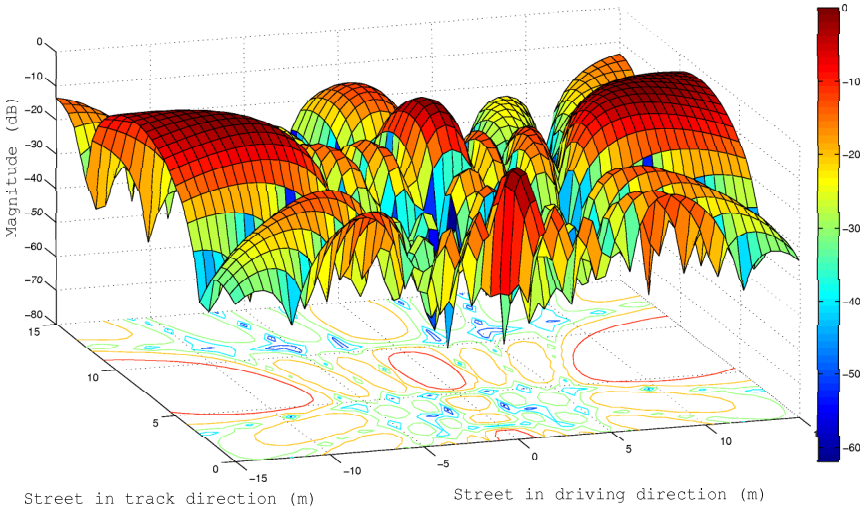

Figure 32: *The response due to the grid array*

Chapter 5

Enhancement of the first solution

After analysing the beamformer and the acoustics situation the first solution, especially the vehicle and the direction detector, has been enhanced.

5.1 The adopted solution

5.1.1 The hardware

In general the beamformer system build up on a box (figure 1) with an exchangeable front plate including the bores for the microphones (the bores of course varries depending on the array design). The microphones itselves are fixed additionally inside the box and can be connected by opening the backside of the case.

The adopted hardware in summary:

- The array design (the spatial microphone arrangement): an crossarray in fact used as two linear arrays (The array is rotated with -45° against the gantry → figure 3)
- The microphone type: PCB Model 130D20 with BNC connector and constant current exitation, free field frequency response
- The measurement interface: National Instruments Interface NI 7422 with synchronisation
- The acquisition data format: TDMS, a streamable data format of NI Labview

5.1.2 The measurement data

Because it is not possible to simulate wrongway drivers on a highway, it has been looking for a most similar situation. This situation is found on a clearway, whereas the speed is higher than on normal streets and additional there are 2 tracks aviable in both directions. The measurements are done on an Austrian freeway.

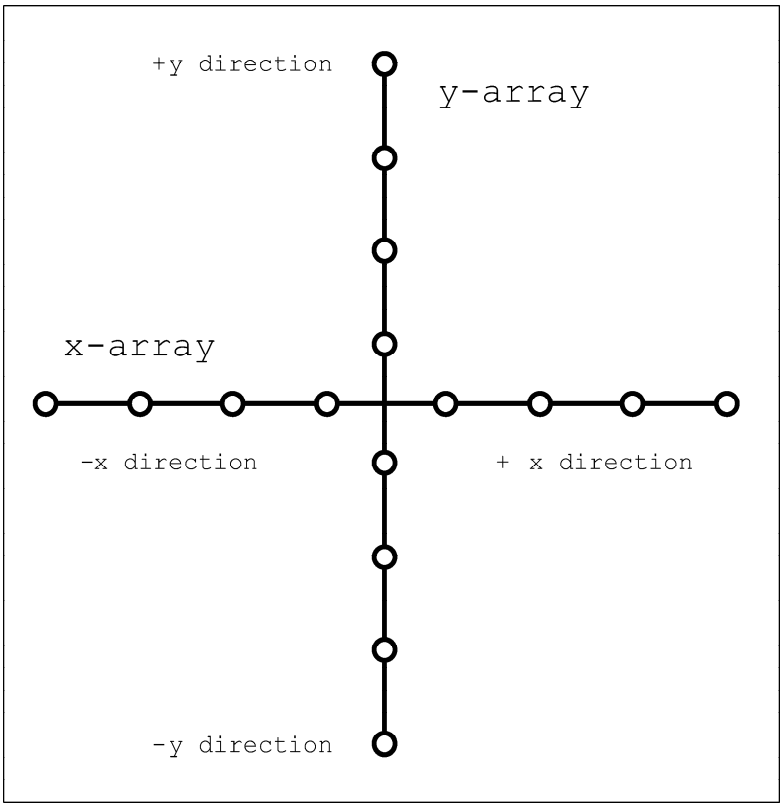

Figure 33: *The two line arrays with the direction defintions*

5.1.3 The adopted software solution

As mentioned in section 1.3 the basic build up of the system / software is a spatial filter or beamformer followed by a car and a directional detector.
The software itself is sectioned into 3 parts:

Chunk data:

The main task of the first program part is the preparation of the data, means loading the „tdms-file“and fragment the gotten mat-files into chunks of data, whose are named by the particular timesection.

Calc power:

The part includes the spatial filter calculations, thus after temporal prefiltering and the calculations of the steering delays, the sum and delay algorithm is applied on every chunk of data. The actual implementation is done by writing the files in a matrix, whereas the rows are the channels and the position in the column is due to their particular delay (dependent on the steering delay and the position of the sensor). Then the sum across the rows is applied, which relates to the steered signal for the particluar steering delay.[1]. Finally the power / rms value is calculated across the data block and thus one value, representing the steering position in a certain time range, is gotten. Due to the usage of 2 independend line arrays the algorithms has to be applied on every line array.

The car/direction analysis:

Before detecting the direction of the car (the main issue) the detecion of the car is essential. This is done by taking the beamformer signal. Temporal filtering is applied to the signal and within a search after maximum values the vehicles can be located. As can be seen in the figure 34 the peaks in the function relates to car positions.

The direction detector involves of course the detection of a car. With the knowledge of the car positions or more exact the time of the car crossing the steering information is taken around this time. To figure out the direction of the vehicle the spatial signal is taken and thus prepared for the direction detection, a line detection (see figure 35).

5.2 Enhancement of the system

5.2.1 Detection theory

Before enhancing the detectors, a few theoretical facts and quantities should be clarified. Basically, the two detectors are 2 Theses Detectors and the theory which handles whith such classificators is a part of the statistical signal processing [8] and is called detection theory. In fact the task of detecting is to decide between two states whereas the 2 states are most of the time „signal“and „no-signal “or just noise.

[1] Another way of realisation could be the convolution with a delayed impulse → relates to a filtercoefficient

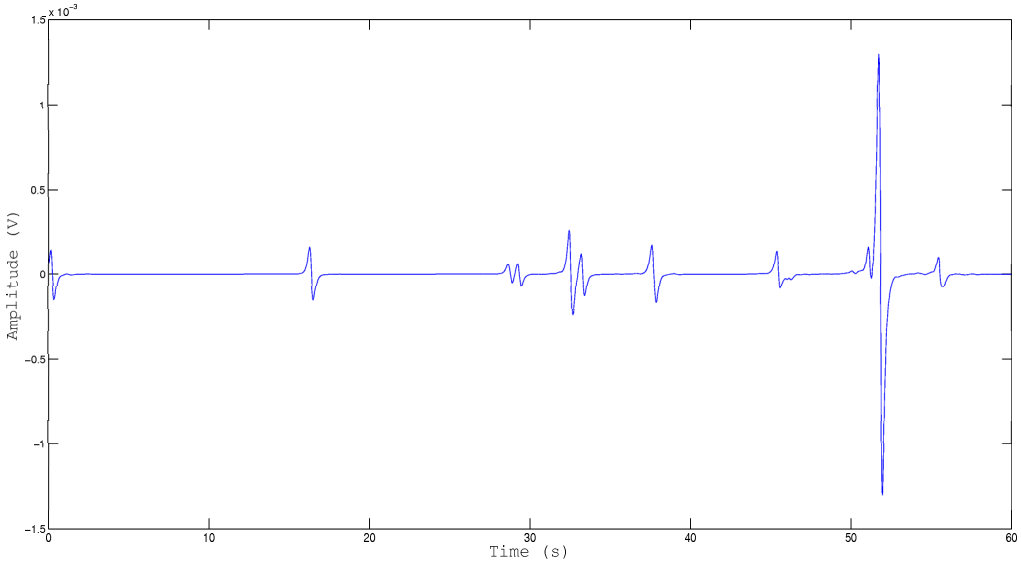

Figure 34: *car detection function*

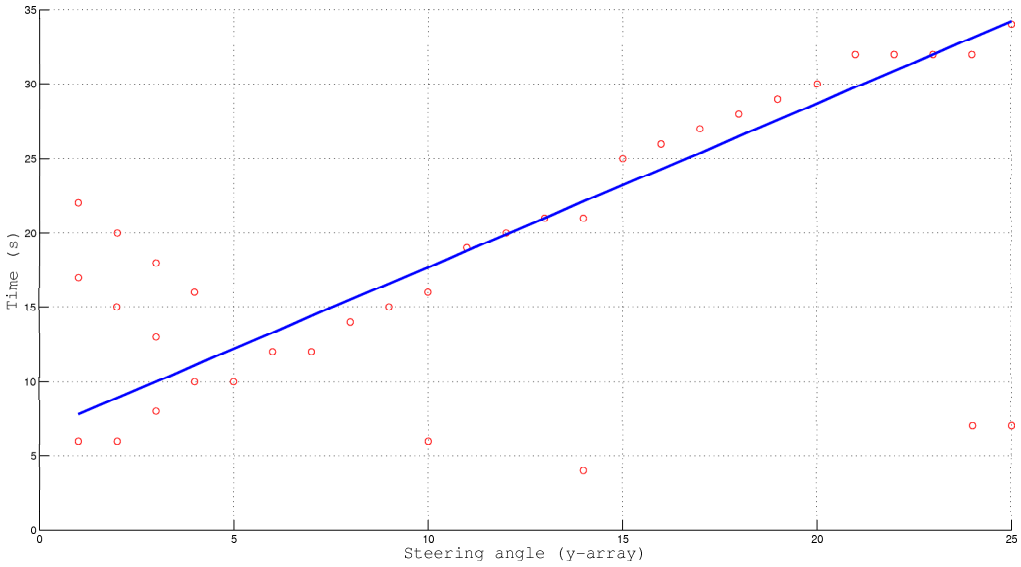

Figure 35: *The diagramm shows the scattered direction-time analyis result by the x-array (across the street in driving direction)*

The basic build up of a detector can be described by two steps (whereas is its sometimes difficult to seperate):

1. Preprocessing: includes any kind of filtering or transformation
2. Feature reduction: includes the threshold(s) to reduce the data to the desired features

Preprocessing

By dropping back to the two states it has to be said the „statelevels“liables some randomness. This randomness is in detection theory modeld by a specific probability function (or frequency distribution), whereas it is common to choose a gaussian probability distribution for any state. Thus the states having there means and a specific variance. Depending on the distance between the means and the variance overlapping will occure. At this point the preprocessing comes into play to enhance the distance between the statemeans and by reducing the variance of the states. To benchmark the preproessing the deflection coefficient d (deflection coefficient d, see figure 36)can be used or indirectly the ROC (see threshold investigation).

It should be emphased that the preprocessor is the most important part of a detectorsystem and without a reasonably prefiltering the treshold of the feature reduction can't be set optimally.

Feature Reduction

After preprocessing, the feature reduction startes with setting one or even a few thresholds. As can be seen in figure 36 depending on the threshold level and the overlapping of the probability functions four decision states can happen:

	Car detected	**Car not-detected**
Car present	True Positiv (Hit)	False Negativ (Miss)
Car not present	False Positiv (False Alarm)	True Negativ (Correct Rejection)

Table 5.1: *The confusion matrix, or in other words the four possible decision states concerning the car detection*

Threshold investigation

By taking into acount the area under the probability function curves it turns out that with increase of the threshold t the detection probability (TP) decreases with the false alarm probability (FP) and vice versa. With this probability function the ROC space with the ROC curve is close conntected. ROC means as mentioned before Receiver Operating Characteristic and it shows the values for the TP probability and the false alarm probability depending on the treshold value. The ROC serves therefore a possibility to benchmark the preprocessing stage of the detection system and shows the optimal threshold by the left upper point in the graph (see figure 36). By changing the threshold (and observing the behaviour of the TP,FP in the

probablity figure) a curve can be sketch with the start in 100,100% and by ending up with t→ infinity in TP and FP = 0 %. This curve describes the complete detector and not like with just one false alarm rate and one hit rate the threshold setting.

5.2.2 Problems with the car detector

The main demand of the car detector is to detect cars with on the one hand high probability of hits and otherwise with low probabiltiy of false alarms. The ideal case would be of course 100 % detection probability and thus 0 false alarms what is just possible with big signal distance between noise and the signal occuring with a car crossing. This is not the case with our system and therefore the derivation of the threshold level between the two stages is the challenge. The term problem with the car detector is concerning the adopted detector maybe overstated because it works well but it should be revealed that there are other possibilities, or car detection function as well.

5.2.3 Enhancement of the car detector

The enhancement of the car detector is done by considering a similar problem in a different area, the Music Information Retrivial (MIR). The Music Information Retrivial deals as the name tells with retriving data from music. The more particular subsection of MIR [9] which is similar to the car detection problem is the „beat or oneset“detection. In the work just 2 different methods of „oneset“detections have been compared, whereas the methodes having just an impact on the preprocessing the feature reduction stays the same. The 2 methodes [9]:

1. The already used „energy base detection“narrowband - band pass filtered signal
2. The high frequency content method (HFC - squared weighted) with broadband signals

To show the difference in the methodes just a shortcut of the signal is used. Figure 37 a) shows the signal once original and second times unfilterd. The plot below shows the gradient of the car detection function based on the different methodes.

Energy based detection

The energy based detection is subordinated to the temporal features reduction methodes, and is calculated by:

$$E(n) = \frac{1}{N} \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} x(n+m)^2 \tag{5.1}$$

This method was already used in the first implemention, the final version differs just in the postprocessing and the threshold (the signal is still bandpass filtered).

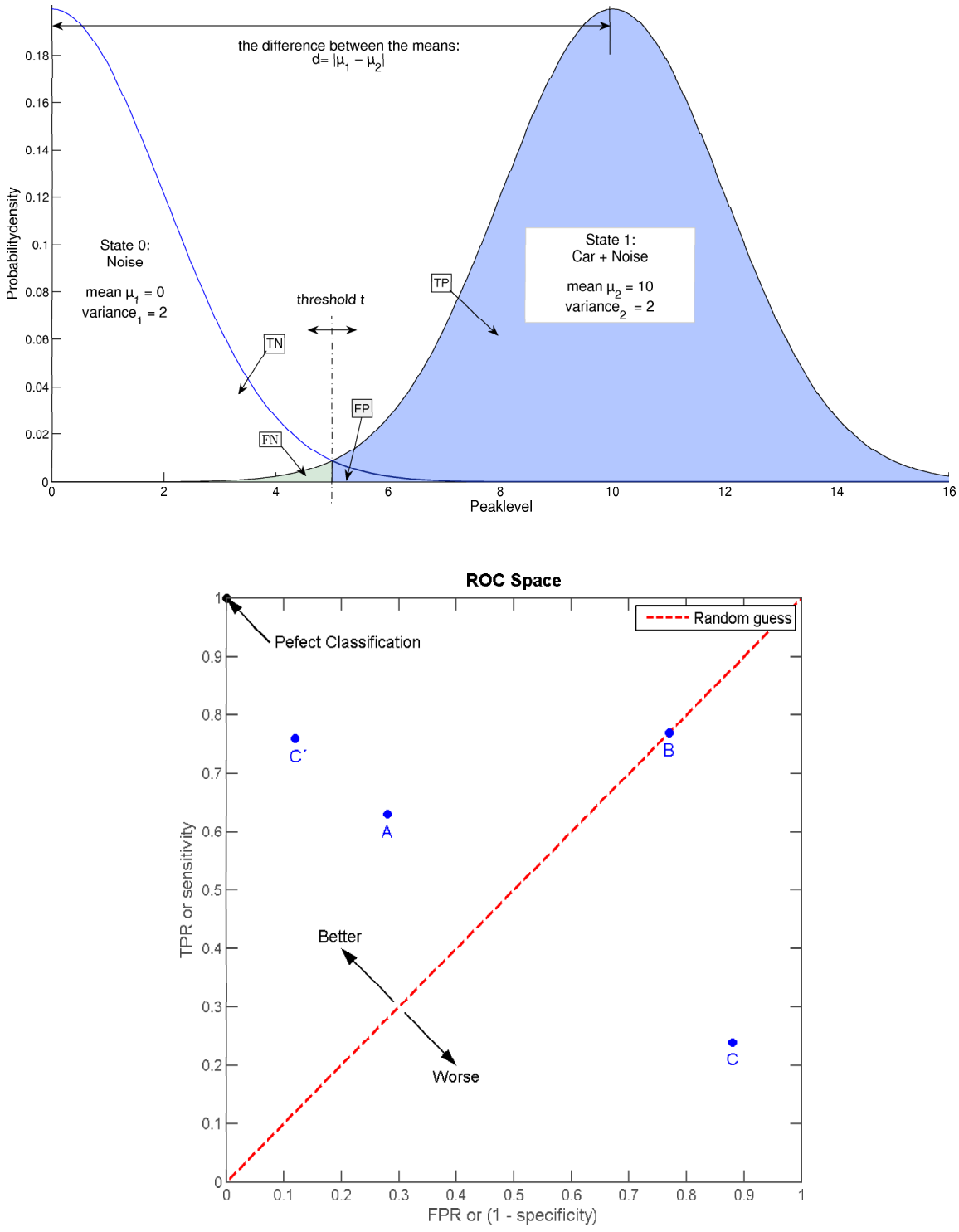

Figure 36: *Above, the two states with its probability distribution (defined by mean and variance) and the 4 decision areas (TP,FP,TN,FN). The decision areas of course variies concerning the threshold, the difference d of the means and the variances. Below, the ROC Space with 4 different points representing threshold values.*

High frequency content method (squared weighted)

The HFC method proposed by masri [9] is one of the spectral feature extraction methods and is defined by, whereas W_k can be defined by $|k|^2$ therefore squared weighting of the spectra:

$$\tilde{E}(n)\frac{1}{N}\sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} W_k \left|X_k(n)\right|^2 \tag{5.2}$$

This method has just considered theoretical to emphase a different possibility of realising the car detection. Therefore the method has not been realised within this project. To emphasis the performance of the detector functions the gradient calculation is done (concering the positiv threshold just the positiv peaks will be accounted). Because as mentioned before the gradient is used to localised the cars. By plotting the narrowband version (energy based method) it results, that the peak mean is lower and the variance is larger than with the broadband method (HFC). On the other hand the noise variance with the HFC is worse than with the energy based method. For that reason the best method has to be found inbetween this methods. This leads to the applied car detection function with smoother filtering an thus a broader (in the frequency sense) signal.

5.2.4 Evaluation of the car detectors

Before doing the evaluation the files had to be annotaded. Thereby the time of the car crossings has been noted by observing the video signal which has been recorded synchronized to the audio signals (the work is done with the audio sequencer software wavelab). In conclusion a matlab file has been used to compare the algorithmic results to the annotation. An important point due to the evaluation of the car detection function is that the non-detections (True Negatives) can't be evaluated.

Again with the evaluation the energy based method used before is compared to the smoother filtered version (of course with different threshold due to the different filtering) done within this work. The results for one measurement are:

	adopted system	**enhanced system**
# of cars detected	206	186
# of cars in the Transcription	184	184
Hits TP	180	181
Misses FN	4	3
Correct Rejections TN	0	0
False Alarms FP	26	5

Table 5.2: *The confusion matrix for both detectorsystem*

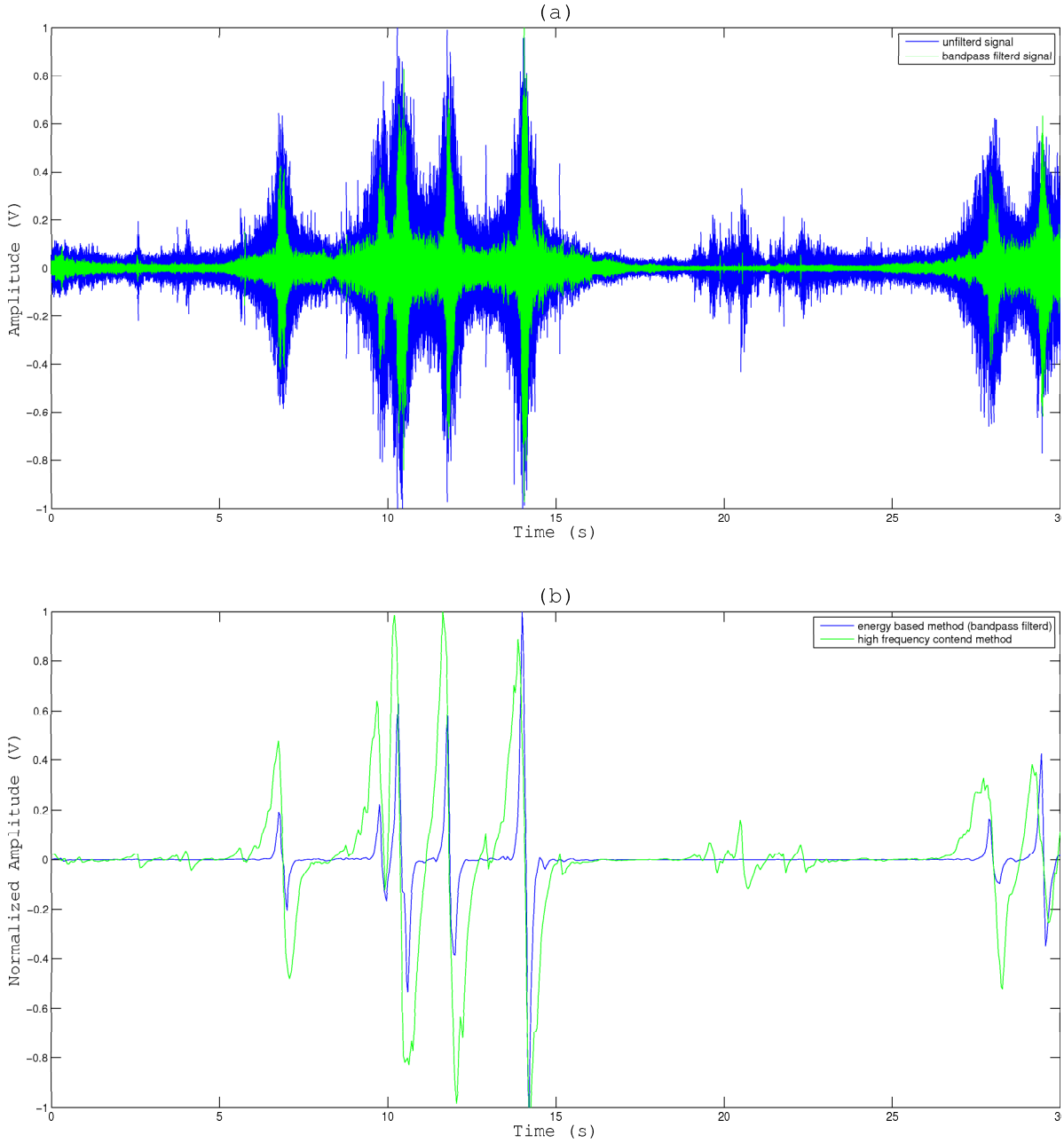

Figure 37: *a) shows the original signal and the bandpass - filtered one. b) the 2 detection functions.*

In the theory part the calculation of the ROC for detector-benchmarking was explained. But within a single set or a single confunsion matrix (for just one treshold) its common to express the performance just by single rates:

- The **True Positve Rate** (Recall or Sensitivy) is the proportion of positiv cases that were correctly identified:
$$TPR = TP/(TP + FN)$$
- The **False Negative Rate** is the proportion of positiv cases that were incorrectly classified as negative
$$FNR = FN/(TP + FN)$$
- The **True Negative Rate** (Specifity) is the proportion of negative cases that were classified correctly
$$TNR = TN/(FP + TN)$$
- The **False Positive Rate** (Fall-out) is the proportion of negative cases that were incorrectly classified as positive
$$FPR = FP/(FP + TN)$$
- The **Accuracy** is the proportion of the total number of predictions that were corrected:
$$A = (TP + TN)/(P + N)$$
- The **Precision** is the proportion of the predicted positive cases that were corrected:
$$PR = TP/(TP + TN)$$

The evaluation for the clearway:

	adopted system	**enhanced system**
TPR	0.97	0.98
FNR	0.021	0.0163
TNR	/	/
FPR	/	/
A	/	/
P	/	/

The results illustrates the enhancement of the car detection function by the increase of the TPR and the FNR. The further rates cannot be calculated or more precise delivers no valid values due to the missing of the "'True Negative"'.

5.2.5 Problems with the direction/speed detector

By observing figure 35 the slope representing the direction of the car can be seen. In the case of a few cars, whose crossing in different directions the analysis with regression lines become much more complicated. With more then one car, more than one regression line is required but with unkown allocation of the points [2]. The problem can be seen in the left lower corner (the points in the left corner) of figure 35 where the situation of 2 crossing cars occurs.

5.2.6 Enhancement of the direction detector

Regressionline

The first solution to solve the problem with more than one line (car) has done by a special two line regression. Therefore the points had been allocated to the two lines and furthermore applying a regression algorithm to both lines. But this works, due to the allocation mechanism, with just cars whose crossing in the middle and not with the general case like in the figure 36.

Hough-algorithm

It could be shown that the problem of direction detection is a line detection problem, which is well known in image processing. This leads to the simplest of the algorithms the hough algorithm with the including hough transformation:

> *To sum up, the hough algorithm is a stable methode to detect lines (or even other shapes) in a "binary - gradient"' image or video. Binary - gradient means the calculation of edge detecion (gradient) and a monochromatic transform (black/white image) before doing the hough transform.*

As mentioned in the citation the hough algorithm is seperated into 3 parts:

1. The edge detection (canny, sobel)
2. The monochromatic transform
3. The hough transform

It now appears because of the two picture transforms there are a few static thresholds in the algorithm. The adaption of this thresholds is done empirical and turned out to be the real challenge of the algorithm. Since the main part of the algorithm is the transform, it has to be explained in more detail.

[2] It has been tried to use correlation to figure out the allocation of the points but there had been no significant correlation

The hough transform:

The hough transform is in general a parameter transform which maps a point (especially an edge-point) into the hough - space. By considering such an edge-point [3] of a line and looking for lines which passes trough the point a function on the line parameters can be drawn (But in general the line is described in another way than just with the incline and the offset → the hessesche normalform). By doing this on all the edge points there will be a crossing between the parameter functions, this crossing leads exactly to the parameters whose defining the line defining the edge-line.

[3]The transform is done on all points on an edge

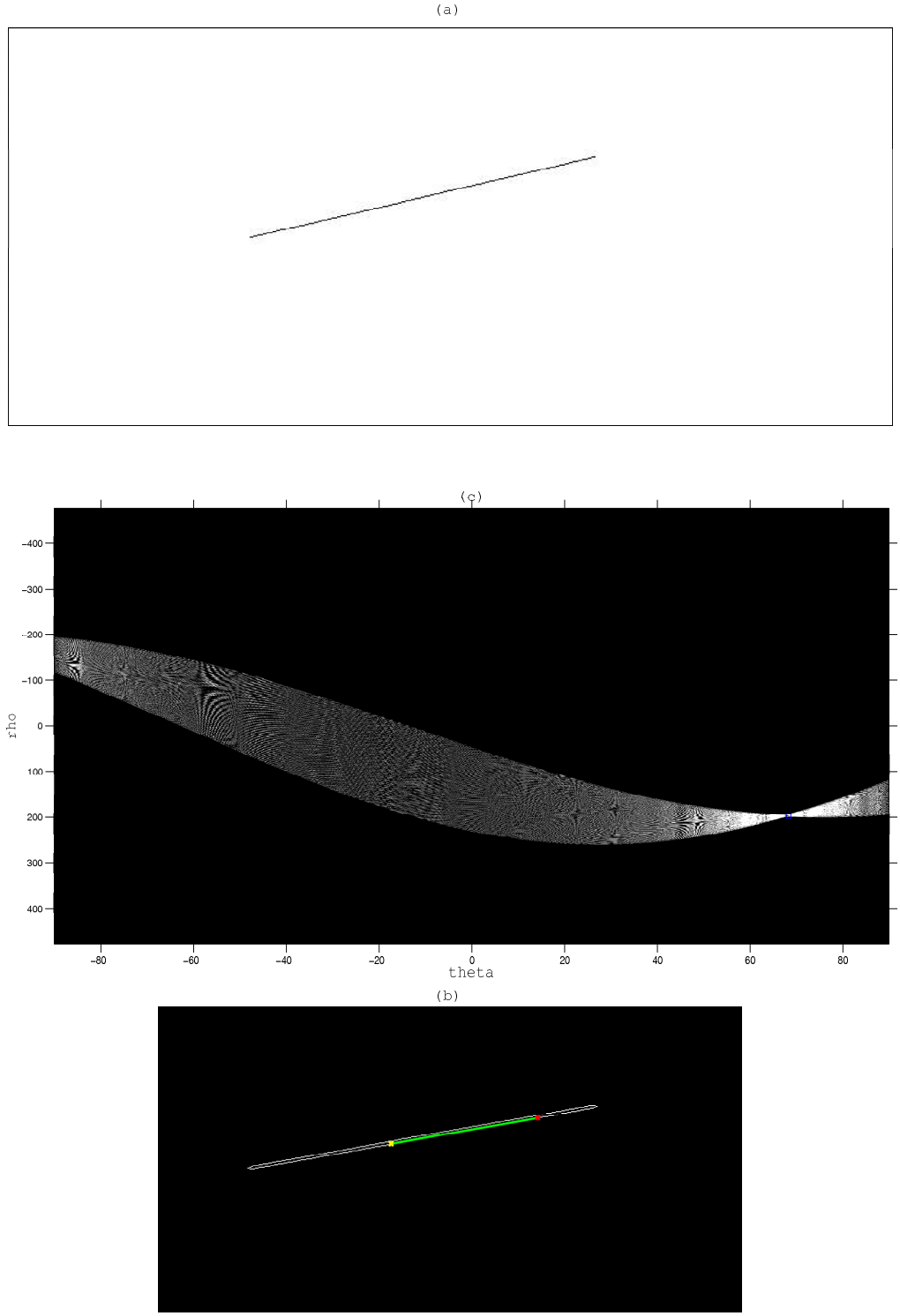

Figure 38: *a) shows the orgin-picture, b) shows the line detection applied to the orgin picture and c) illustrates the houghspace*

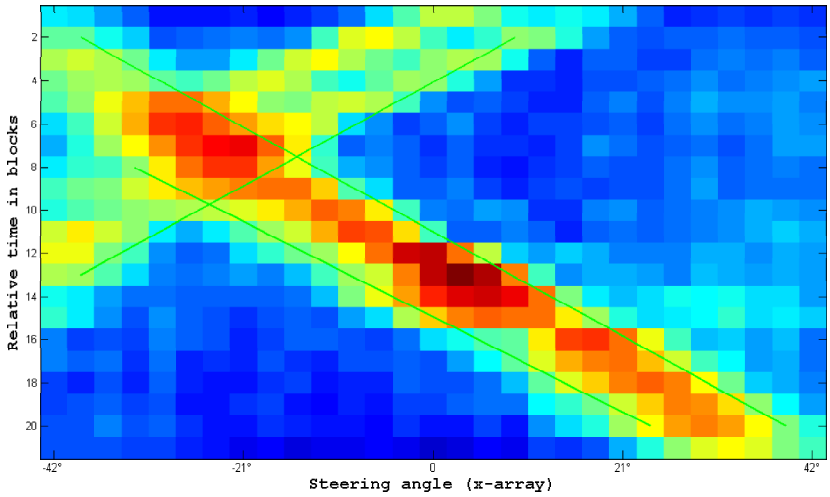

Figure 39: *The plot illustrates the line detection solved by the hough algorithm*

The results by the hough algorithm

As brought up in the „problems of the direction detector“the problem with the first solution was the detection of more than one lines, at the same time. With the hough algorithm the problem can be solved as can be seen in figure 39. The figure shows the direction - time analysis around a car detection point, whereas the green lines represent the lines whose have been derived by the hough algorithm.

5.2.7 Evaluation of the direction detector

The evaluation of the direction detector is done alike with the car detector. It should be mentioned that the speed connected with the direction has not been evaluated because it is just an additional feature and not that important. The confusion matrix of the direction detector for Sieggraben results in:

	adopted system	**enhanced system**
# of cars detected	206	186
# of cars in the Transcription	184	184
Hits TP	77	92
Misses FN	19	4
Correct Rejections TN	79	84
False Alarms FP	5	1

With this confusion matrix the rates results in:

	adopted system	**enhanced system**
TPR	0.80	0.97
FNR	0.197	0.041
TNR	0.94	0.98
FPR	0.059	0.011
A	0.87	0.97
P	0.939	0.98

The evaluation of the driving detection was done in contrast to the car detection with the „True Negativ"cases thus all the rates are available and deliver better results than with the first solution.

Results and Suggestions

By considering the results of both detectors, the most critical events are the „miss"and the „false alarms"(whose are positionend in the side diagonale of the confusion matrix). But with this system the really dangerous one is the „miss"of a wrong way driver. Because if it is a false alarm (FP) the second instances comes into role (can be a person or a backup system) and can disable the alarm after control. But with a „miss"the wrong way driver will not be observed until the next gantry in a few kilometers (what can be too late).

Additionally it would make sense to combine the probabilities of both detectors, thus to express for example the probability of true detected directions if the cars are true detected as well. But this is of course not done by the product of the probabilities because the detector systems work not statistically independent and thus the „conditional"probability has to used (bayes equation).

Finally it should be mentioned even with the enhanced system there are cases whose can not be detected in a right way (40). As a direct consequence it should said, thats due to the bad beamformer performance and thus it turns finally out that:

The development of a well working detector is basically settled by the work of the preliminary filter stage

This statement is of course valid for spatial and temporal filters as well.

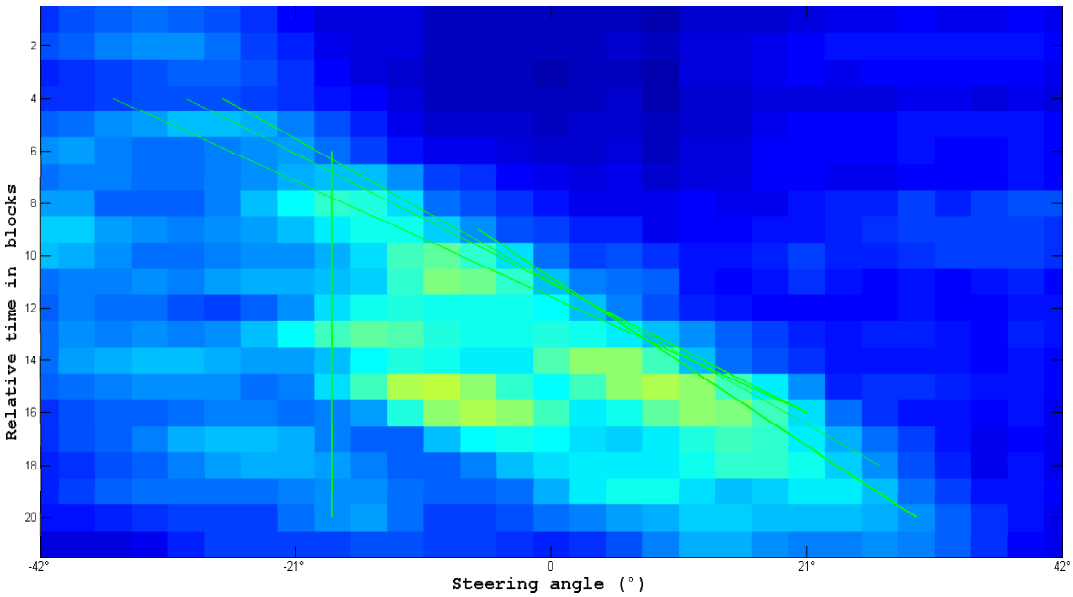

Figure 40: *The figure illustrates a problematic case for the hough detection*

Chapter 6

Suggestions and Endcomments

In the end of this thesis some suggestions and endcomments are essential. In general it can be said: The sum and delay beamformer algorithm is in fact the easiest one, but with beamformer it is always dependent on the acoustic situation as well. Whereas like in our case with the distortions due to the projection of the system to the street the solution becomes really complicated. Altought the greatest challenge would be to handle with reverberation as well → just first reflections had been occured.

Additionally the analysis of those situations by the measurement data is hardly possible and can just be simulated. Therefore the simulation tool is practicable but offers of course not the exact beampattern, especially with the boundary layer the behaviour has to be different (the boundary layer assumption can be added in the simulation as well).

Concerning the car detection and direction detection problem it could be seen that those problems are not unknown other scientists having similar problems in different areas. Thus in the solving process its really helpful to „sniff“into different areas as well.

Another point has to be said about the array design or finding the optimal spatial sampling positions. In the spatial signal processing their are a few methods to develop the optimal array design but most of them are just valid for the narrowband case and therefore for the broadband version to complex. But it could been seen that within a good array the performance of a beamformer can be increased a lot. Thus I think in most of the cases its not useful to handle with adaptive beamformer but rather use a easier algorithm and figure out the best array design.

Chapter 7

Bibliography

[1] Priebsch H-H.: Akustik für Motor und Fahrzeug, VO-Skript, TU Graz, SS 2007

[2] Bruel & Kjaer.: Beamforming Review No1, B&K, 2004

[3] Benesty J.; Chen J.; Huang Y.: Microphone Array Signal Processing.1.Auflage.Springer,Berlin 2008 pp 39-65

[4] Mc Cowan I.: Microphone Arrays: A tutorial, Phd Thesis, Queensland University of Technology, Australia 2001

[5] Mohammad S. K.; Roger M. G.; Dixon R.: Implementation of non-uniform sampling for alias-free processing in digital control,Controls Systems Group, Department of Electronic and Electrical Engineering, Loughborough University, 2008

[6] Herbordt H.; Kellermann W.: Beamforming for Audio Signal Acquisition, A short course for Temic Speech Dialogue Systems, University of Erlangen-Nuremberg, March 2004

[7] Tashev I.; Malvar S. H.: A new beamformer design algorithm for microphone arrays, Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA

[8] Brander M.: Statistical signal processing, Vorlesungskriptum, TU Graz, SS 2008

[9] Bello J.P.; Daudet L.; Abdallah S.; Duxbury C.; Davie M.; Sandler M.B.; A tutorial on onset detection in music signals. IEEE Transactions on Speech and Audio Processing. Scheduled for publication on September, 2005.

[10] Silverman F.H.; DiBiase J.; Brandstein S. M.; Robust Localistion in Reverberant Rooms, Brown University, Harward University, 2001

[11] Knapp C. H.; Carter. G. C.; The Generalized Correlation Method for Estimation of Time Delay, IEEE, 1976

[12] Wien konkret - Verkehr - Auto Autobahnen - Geisterfahrer Autobahn .: http://www.wien-konkret.at/verkehr/auto/geisterfahrer/ ,2008

[13] Goodwin M.M.: Frequency independent beamforming, Acoustics Research, AT&T Bell Laboratories, and EECS, University of California at Berklay 1993.

List of Figures