

Slovak University of Technology in Bratislava
Faculty of Electrical Engineering and Information Technology
Department of Telecommunications

A new methodology for
the design and the quality control
of the acoustic feedback
of control elements in cars

Dissertation
by

Dipl.-Ing. (FH) Alexander S. Treiber

October 2005 – May 2010

Beta version, not to be distributed

Supervisor: doc. Ing. Gregor Rozinaj, PhD.
Co-Supervisor: Prof. Dipl.-Phys. Gerhard Gruhler

Contents

List of Figures	viii
List of Tabela	ix
List of Symbols	xi
1 Introduction	1
1.1 Purpose of this Study	1
1.2 Acoustics in the Human Machine Interface	1
1.3 Objective of this Work	5
1.4 Specification of Chapters	5
2 Introduction to Sound and it's Measurement Methods	7
2.1 Nature of Sound	7
2.1.1 Sound Pressure	7
2.1.2 Spectrum	9
2.2 Processing by the sense of hearing	12
2.2.1 Preprocessing	12
2.2.2 Perception	14
2.3 Common methods for subjective analysis	18
2.3.1 Paired Comparison	18
2.3.2 Semantic Differential	18
2.3.3 Adjustment	18
2.3.4 Magnitude Estimation	18
2.3.5 Adaptive Methods	19
3 State of Research	20
3.1 Analysis and Synthesis of Transient Noise	20
3.1.1 Musical Applications	20
3.1.2 Virtual Reality and Games	21
3.1.3 Sports	21
3.1.4 Sound Quality	22
3.2 Applications of acoustics in human machine interfaces, cossmodality studies .	23
3.2.1 Technical University Dresden - Altinsoy et al.	23
3.2.2 University of Glasgow - Brewster	23
3.2.3 NASA Ames Research Center - Begault	23
3.2.4 University of Oldenburg - Töpken	24
3.2.5 Norman	24

3.2.6	Technical University Munich - Fastl et al	25
3.2.7	Heilbronn University - Reisinger	25
3.2.8	AIDE Project	26
4	Goals and Approach of this Work	27
4.1	Current Methods of Specifying and Measuring Transient Sounds	27
4.2	Discussion of the Limitations and Shortcomings of the Currently Used Methods	27
4.2.1	Specification of Maximum Sound Pressure	27
4.2.2	Practical Implementation of Subjective Evaluation	28
4.2.3	Highly Subjective Target Sound Specification	29
4.3	Proposed Improvements	29
4.3.1	Introduction of Spectral and Temporal Properties	29
4.3.2	Introduction of a Acoustic Rapid Prototyping Methodology for Sound Design	31
4.4	Theoretical and Practical Approach	31
4.4.1	Iterative Examination of Sound Properties in Subjective Trials	31
4.4.2	Isolation of Properties Relevant for Subjective Perception	32
4.4.3	Deviation of a Synthesis Model for Transient Sounds	32
4.4.4	Deviation of an Analysis Algorithm	32
4.4.5	Experimental Evaluation of Analysis Algorithm in a Subjective Trial	32
5	Preparation	33
5.1	Measurement Equipment	33
5.1.1	Robotic Acutation Unit	33
5.1.2	Anechoic Test Box	35
5.1.3	Recording, Signal Conditioning and Data Extraction	36
5.1.4	Analysis Tools	37
5.2	Description of the acoustic feedback signal	37
5.2.1	Amplitude	38
5.2.2	Spectral Composition	39
5.2.3	Evaluation of Real Control Elements	39
5.2.4	High-Speed Video	41
5.3	Interactive Acoustic Simulator	43
5.3.1	First Version	43
5.3.2	Improved Version	43
5.4	Listening Room	47
5.4.1	Layout	48
5.4.2	Simulation and Choice of Materials	48
5.4.3	Acoustical Properties	49
6	Synthesizer	50
6.1	Basic principle	50
6.2	Modification of Amplitude Envelope	50
6.3	Introduction of Wavetable Oscillator	51
6.4	Control Parameters	53
6.5	Synthesis Tool KLKR	54
6.5.1	Oscillator	54
6.5.2	build envelope.m	55

6.5.3	Phase correction of the oscillator	55
7	Subjective Experiements	56
7.1	Paired Comparisons of Recorded Sounds	56
7.2	Paired Comparisons of Loudness Normalized Synthetic Sounds	58
7.3	Rating of Synthesized Sounds	60
7.4	Search for JNDs and Reaction Time Measurements	62
7.4.1	Design of Experiment	63
7.4.2	Results	67
8	Proposed New Method for the Prediction of the Acceptance of Acoustic Feedback	72
8.1	Outline	72
8.2	Analysis Program "proBand"	72
8.2.1	proband.m	73
8.2.2	Tools	75
8.3	Correlation	76
8.4	Judgement / Selection	76
8.5	Experimental Validation of the Predictor	78
8.5.1	Design of Experiment	78
8.5.2	Analysis of the Results	81
8.5.3	Conclusion	86
9	Results	87
9.1	Method for Acoustic Rapid Prototyping	87
9.2	New Algorithm for Prediction of Acoustic Acceptance	88
9.3	New Proof for the Effects of Acoustic Feedback	88
10	Conclusion	89

A Data Sheets	91
B Source Code	109
C Raw Data from Subjective Trials	110
Glossary	111
Literature	113

List of Figures

1.1	Dashboard of a 1970 VW Beetle	2
1.2	Dashboard of a BMW E36	3
1.3	Dashboard of a BMW E65	4
2.1	Human listening plane	8
2.2	A-, B-, C- and D-weighting curves	8
2.3	Frequency analysis using a bank of parallel bandpass filters	10
2.4	Explanation of frequency analysis plots	11
2.5	Schematic drawing of the ear	12
2.6	Frequency analysis on the basilar membrane	14
2.7	Nonlinear perception of pitch	15
2.8	Critical bands and masking	16
2.9	Weighting function for sharpness	17
3.1	Haptic simulator	25
3.2	Prototype of AIDE interface	26
4.1	wahrnehmung0	29
4.2	wahrnehmung1	30
4.3	wahrnehmung2	30
5.1	First Version of the Actuation Robot	34
5.2	Final Version of the actuation robot	35
5.3	Connection Terminal Outside	36
5.4	Connection Terminal Inside	36
5.5	Typical micro switch on PCB	38
5.6	Typical rotary encoder PCB	38
5.7	Typical acoustic feedback signal ("click") in the time domain	38
5.8	Typical acoustic feedback signal ("click") in a spectrogramm	39
5.9	Extremely loud sample	40
5.10	Extremely quiet sample	40
5.11	The frame shows the state immedeately before the movement of the spring	41
5.12	The spring is in motion to the right	41
5.13	The spring has hit it's impact	41
5.14	No motion of the spring observeable 3 ms after the first frame	42
5.15	The Yamaha A5000 sampler which is used for playback in paired comparisons	43
5.16	Flowchart of the Simulator	44
5.17	Generation of trigger signals from the hi-resolution encoder	45
5.18	Response of the acoustic simulator to a trigger signal	46

5.19	Duration of one loop of the microcontroller	47
5.20	Layout of the jury test room	48
5.21	Reverberation time of the listening room	49
6.1	logattack	51
6.2	wavetable	52
6.3	wavetableclick	53
6.4	Dull Spectrum	54
6.5	Bright Spectrum	54
6.6	phasecorr	55
7.1	All sounds used in the first series of tests	56
7.2	Results of the first jury test	57
7.3	All sounds used in the second series of tests	59
7.4	Results of the second jury test	60
7.5	j3decay	61
7.6	j3spectra	62
7.7	jndprocedure	63
7.8	j4s2interface	64
7.9	updown	65
7.10	j4s3interface	66
7.11	j4s2soundsettings	68
7.12	j4s2spectra	68
7.13	j4s3fastest	69
7.14	j4s3accurate	70
7.15	j4s3accuracymean	70
8.1	envin	74
8.2	noisefloor	75
8.3	User interface for Sound design in the final experiment	78
8.4	Example for the Distribution of Correlation Coefficients	79
8.5	Interface in the final test stage	80
8.6	Individually optimal spectra	82
8.7	Individually optimal Damping Coefficients	83
8.8	Individually desired peak amplitudes	84
8.9	3D Plot of Acceptance vs. Correlation	85
8.10	Influence of the peak amplitude	86
9.1	Rapid Prototyping Methodology	87

List of Tables

7.1	j4s3sequence	67
7.2	j4s3sum	71
8.1	Table in j6	81

List of Symbols

α	Absorbtion coefficient
λ	Wavelength
c	speed of sound
f_s	sampling frequency
L_p	sound pressure level
T_{60}	Reverberation time
K	Amplification factor

Chapter 1

Introduction

This introductory chapter gives a total overview of this work, responding in detail to the actual task of the industry and consequently providing an insight in the approach to solving this task.

1.1 Purpose of this Study

Due to the complexity of modern car's on-board systems car manufacturers tend to use menu based user interfaces. These interfaces offer a relatively low number of control elements for a large number of functions. Optimization of both haptical and acoustical feedback of the remaining control elements can improve both acceptance by the user and security of operation. Furthermore, there is a trend towards the use of touch-screens and capacitive buttons which offer no inherent auditory and tactile feedback. However, adding synthetic feedback to this kinds of input devices improves usability.

Since cars nowadays are no longer sold simply as technical but as lifestyle products it is crucial for the success that the potential customer perceives every single aspect of the car to be valuable. Obvious examples are exterior design and material quality. Since the user interface of the car can be judged even before a test drive the haptical and acoustical feeling of buttons and switches is a key aspect.

The aim of this study is to fulfil the industry's need for a reliable way of specifying acoustic feedback signals as well as test methods. This is done by recording the acoustic feedback of current control elements and benchmarking the recorded sounds with test persons. From this data a calculation model has to be derived.

1.2 Acoustics in the Human Machine Interface

There are multiple applications for acoustics in the human machine interface (HMI). One is the aspect of warning signals which are generated using electronics and played back to the driver using loudspeakers; another is speech recognition. Although not widely used at the moment it has potential for useage in future HMIs as progress in both speech recognition

algorithms as well as microphone technology makes this method of input less prone to false detection. Speech synthesis and playback is already common in satellite navigation systems.

This study however focuses on the sound which control elements like buttons and switches radiate upon actuation. This acoustic feedback of switches has two functions. On the one hand it can improve the security of the car by giving more precise feedback of the operation. On the other hand acoustic feedback has to be carefully designed in a way that potential customers at least do not find disturbing and ideally in a way that fits into the design concept of the entire car and is perceived in a positive way. There are usually a different types of control elements in one single panel of a car. A common example is a climate control interface which features:

- Push Buttons for activation and mode selection
- Rotary Switches for temperature selection and fan speed

The car manufacturer might want all control elements on this panel to feel similar.

Security of Operation

While driving a car the driver should be focussed primarily on the road. In cars like the 1970 VW 1302 ("Beetle") as shown in *figure 1.1* very little distraction from this primary task was offered to the driver.



Figure 1.1: Dashboard of a 1970 VW Beetle. Only the radio in the center is not involved in the primary task of driving

Besides from a display for the current speed and the total mileage there are only a few buttons for the lights and a stick switch for the turning indicator. The only items not involved in the driving task are the vase and the radio. Only the latter requires any attention to operate. With only two knobs for tuning and volume operation is very simple. As the complexity of the car increases, the distractions for the driver increase as well. In a modern car the driver is confronted with a wide range of functions:

- Satellite navigation
- Multimedia (radio, CD, DVD, MP3 playback)
- Telecommunication (telephone, internet)
- Climate control
- Comfort features (electric windows, electric seats, ...)
- Car features (settings of dampers, gearbox, traction control, ...)

Traditionally every function of the car had it's individual control and/or display. Hence with increasing functionality the number of displays and controls increased as can be seen for example in the 1990s BMW 3 Series (E36 platform) in *figure 1.2*.



Figure 1.2: Dashboard of a BMW E36. Three displays on the center console, a mobile phone and over 30 control elements are needed to access each function - even without satellite navigation

In this design there are three displays on the center console which inform the driver about the status of the radio, the climate control and give auxiliary information about exterior temperature or fuel consumption. Over 30 buttons and knobs are required to use these features but none of them is actually required for driving. Additionally a mobile phone is installed in the car. The result is described by Norman [1] as the paradox of technology:

”The same technology that simplifies life by providing more functions in each life also complicates life by making the device harder to learn, harder to use. This is the paradox of technology.”

According to [2] the number of controls in a car has risen from just under 30 in the 1950s to around 75 in the late 1990s, displays almost quadrupled from around 10 to around 40 in the same period. In order to simplify the user interface, menu based systems were introduced and are nowadays available throughout the entire model range of BMW (iDrive, see *figure 1.3*), Audi (MMI) and Mercedes (COMMAND).



Figure 1.3: Dashboard of a BMW E65. One raised display close to the driver's line of sight as well as one large control element.

These systems all use a large display and a separate control element or a group of control elements. All systems use a large rotary switch with push function to navigate through the menus and to select functions. Depending on the manufacturer's philosophy there are still dedicated controls available for very common tasks like the volume of the sound system and electric windows.

In any way those systems should provide the driver with the information he needs and let him use the function he desires in a convenient and quick way. Partly this has to be ensured by designing the layout of the user interface according to guidelines explained in the European Statement of Principles on Human Machine Interaction (ESoP). Among other requirements which focus on the actual interaction with the driver, which shall not be discussed here, it is stated that the display has to be mounted highly and the controls have to be easy to reach and *shall not require looking at them*.

To ensure the latter it is helpful that the controls provide the driver with sufficient feedback during operation which he can trust using other senses than his eyes. This can be achieved through a distinctive change of force (for push buttons) or momentum (for rotary switches). This haptic feedback is examined by Reisinger et al [3]. A concurrent acoustic stimulus can amplify the haptic feeling [4] or even dominate it as proofed by Hötting [5]. Finally Kurosu [6] points out that people perform better with equipment which is aesthetically pleasant although this is culturally dependent as Tractinsky [7] discovered. However, the point of aesthetics links the topic of security with the topic of perceived quality. For further information of the mentioned research project please refer to chapter 3.

Perciefed Quality

The acoustic feedback of control elements affects the customer's perception of the entire car's quality subconsciously is the other important aspect. Nowadays cars of different manufacturers within the same class all have similar layout, similar luggage space, similar engines and safety features. Car manufacturers struggle to invent new features to convince the customer of the superiority of their product.

Two key aspects that have not been mentioned this far are the image that is connected to a certain brand and the value the customer gets for his money. A brand's image can be transported through the styling of the products. Value for money depends on all points mentioned above and the build quality of the product, more specifically the perceived quality [8]. The perceived quality can be affected through the choice of materials, clearances, tactile quality of surfaces and sounds. As mentioned above this work deals with the perception of sound quality.

In the past acoustics in the automotive field mainly dealt with engine intake and exhaust noise. With improved dampening of this noise sources road and wind noise became the dominant source. To minimize these noises in the passenger cabin manufacturers use insulation material and in higher classes even double glazing in modern cars.

Furthermore there are sounds that can be heard even when the car's engine is not running. For example there are the sounds of a closing door or of buttons and switches which this work precisely focusses on. These kind of sounds are important to the manufacturers as the potential customer can hear them instantly on his first contact with the car, even before a test drive.

According to the psychologist Peter Glanzmann [9]

People who are focussed on aesthetics decide within milliseconds, whether they buy a car or not.

Hence it is reasonable for car manufacturers to design the sounds of their control elements in a way that customers like and in a way that supports the impression of high build quality.

1.3 Objective of this Work

Purpose of this study is to benchmark the quality of the acoustic feedback of control elements with test subjects and to derive a calculation model which allows the prediction of the acceptance of acoustic feedback signals. This includes the conception and development of suitable equipment for this study. The experience gained from the subjective tests are to be used to develop an objective scale for the acceptance of acoustic feedback sounds. This scale can be used to compare measured signals to pre-defined desired sounds. The tools for target sound design which will be developed in the process can also be used for applications where synthetic acoustic feedback is needed.

1.4 Specification of Chapters

1. This introduction chapter
2. An overview of terms and definitions which are common in the acoustic measurement of sounds and especially in the psychacoustic evaluation. This also includes an introduction into methods common in subjective evaluations.
3. A synopsis of research done in similar fields

4. Formulation of the actual challenges which shall be dealt with in this work including the proposed improvements to the current situation as well as an outline of the steps which will be taken in this project.
5. An overview of the tools which had to be conceived, developed and built as preparatory work at the beginning of this project
6. A summary of subjective experiments which were required to gather sufficient data to formulate a hypothesis to solve the main challenge of the project. Furthermore new insights to the effect of auditory feedback to the performance of the user will be presented.
7. Formulation of the proposed analysis methodology and the proposed approach for specifying target sounds in the future.
8. Experimental validation of the analysis model
9. Summary of the new insights gained from this work
10. Conclusions, discussion of the results and possible further challenges

Chapter 2

Introduction to Sound and it's Measurement Methods

In this chapter the fundamentals of sound and it's propagation as well as the reception of sound by the human sense of hearing are explained. Common psychoacoustic scales which are based on the knowledge of the ear's preprocessing are introduced. The physical properties of the specific signal this work focusses on and common methods for subjective analysis are described.

2.1 Nature of Sound

Sound basically is vibration of particles in a medium (gas, liquid, solid). Sound propagates as a longitudinal pressure wave through the medium. The pressure wave can be reflected or diffracted and can be subject to interference. The propagation speed of sound in air can be calculated as:

$$c = (331.4 + 0.6\theta) \text{ m/s} \quad (2.1)$$

with θ being the temperature in centigrade.

2.1.1 Sound Pressure

The amplitude of the pressure wave describes changes in air pressure which are called sound pressure. Sounds can be described by means of the time-varying sound pressure, $p(t)$. Compared to the atmospheric sound pressure the changes caused by sound sources are extremely small. The unit of sound pressure is the Pascal (Pa). The human ears are capable of picking up sound in the frequency range of 20 - 20000 Hz. The upper limit is decreasing with increasing age. The human ears are most sensitive in the frequency range of 2 - 5 kHz. The auditory threshold for undamaged ears is a sound pressure of 10^{-5} Pa whereas the threshold of pain is reached at 100 Pa.

In order to deal with this broad dynamic range the sound pressure level L_p is usually used to quantify the amplitude of sounds. The sound pressure is related to the sound pressure level by the equation:

$$L_p = 10 \log_{10} \left(\frac{p_{rms}^2}{p_{ref}^2} \right) = 20 \log_{10} \left(\frac{p_{rms}}{p_{ref}} \right) \text{dB} \quad (2.2)$$

with p_{ref} being $2 \cdot 10^{-5}$ Pa. Furthermore the hearing threshold is frequency-dependent as can be seen in *figure 2.1*.

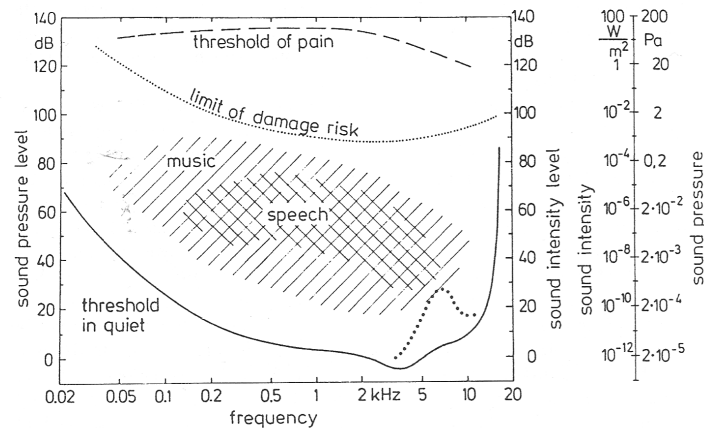


Figure 2.1: Human listening plane, the baseline is the hearing threshold

The sensitivity of the human ear decreases for high and low frequencies. As a result of this a 2500 Hz tone of a certain sound pressure level would be perceived louder than a 250 Hz tone of the same sound pressure level. To compensate this effect weighting filters (see *figure 2.2*) have been introduced, with the A-curve being the one most commonly used.

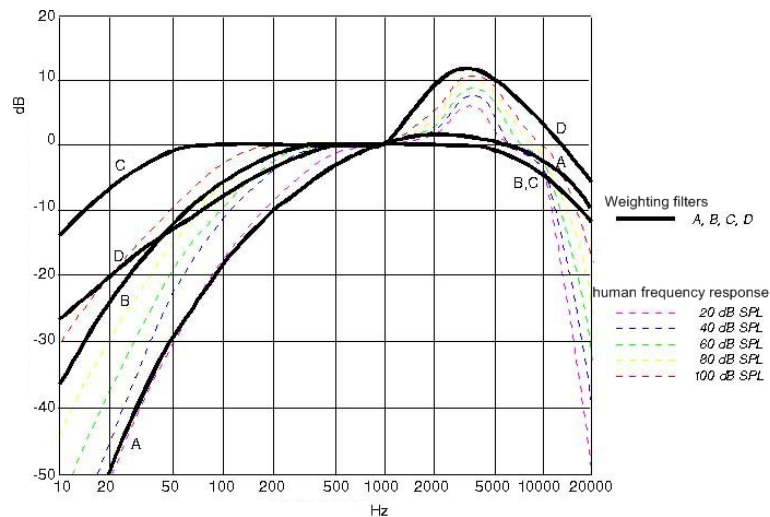


Figure 2.2: A-, B-, C- and D-weighting curves

As can be seen from the figure, also the frequency response of the ear is also not constant

but level-dependent. The change in sound pressure level caused by the weighting filters can be calculated as:

$$K_a(f) = \frac{12200^2 \cdot f^4}{(f^2 + 20.6^2) + (f^2 + 12200^2) \cdot \sqrt{f^2 + 107.7^2} \cdot \sqrt{f^2 + 737.9^2}} \quad (2.3)$$

$$K_b(f) = \frac{12200^2 \cdot f^3}{(f^2 + 20.6^2) \cdot (f^2 + 12200^2) \cdot \sqrt{f^2 + 12200^2}} \quad (2.4)$$

$$K_c(f) = \frac{12200^2 + f^2}{(f^2 + 20.6^2) \cdot (f^2 + 12200^2)} \quad (2.5)$$

The widely used A-curve is only suitable for relatively low sound pressure levels. It can be seen from the unit of the sound pressure level if a weighting curve has been used (i.e. $dB(A)$ for A-weighted sound pressure level).

As mentioned above the sound pressure varies with frequencies of up to several kilohertz. The sound pressure level which is displayed in a sound-level-meter however is obtained by integrating the sound pressure over a relatively long period of time. Three integration times or time constants are defined in DIN EN 61672:

- S (slow): 1000 ms
- F (fast): 125 ms
- I (impulse): 35 ms rise- and 1500 ms decay time

The choice of the time constant depends on the signal which has to be measured. Stationary sounds like the sound emissions of motors are typically measured using the slow setting while for example music or speech should be measured with the F setting due to their fluctuations. The impulse setting responds very quickly and can therefore capture for example impact sounds, due to the long decay time it is possible to read a value despite the short response time.

Sound pressure levels are usually displayed as a two dimensional plot in which the abscissa denotes the time and the ordinate denotes the sound pressure.

2.1.2 Spectrum

The next important property of an acoustic signal is it's spectrum. The sound pressure level of an acoustic signal only contains information about it's amplitude but cannot provide cues to the actual content of the signal. For further information of the sound it is therefore necessary to know which frequencies are contained in the signal. Basically there are two approaches to analyze the frequency components of a signal.

Filters The spectrum of a periodic signal can be examined using a tuneable bandpass filter and a sound level meter. The spectral distribution of sound pressure level can be obtained by successively tuning the filter to different center frequencies. Whereas the steepness of filters in most signal processing applications is specified in dB per decade filters for acoustic measurements are described using db per octave. The bandwidths of the filters are usually

specified in octaves of n -th fractions of octaves. Most common are octave- and 3rd-octave filters. The frequency ratio between the center frequencies of octave band filters is 2:1, the ratio for 3rd-octave 5:4. This also means that the bandwidth of the filters increases for high frequencies as the frequency ratios remain constant.

For nonperiodic signals a bank of bandpass filters as can be seen in *figure 2.3* has to be used.

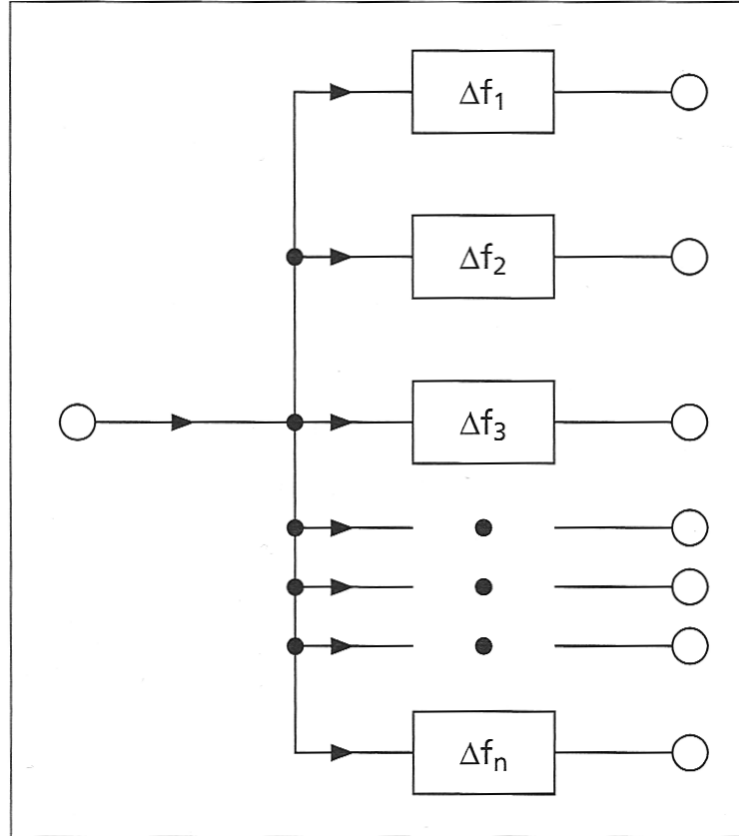


Figure 2.3: Frequency analysis using a bank of parallel bandpass filters

The parallel structure requires simultaneous calculation of the sound pressure level for all bands, however, it allows realtime analysis of signals with time-varying spectral composition.

3rd-octave and octave filters are specified in the DIN EN 61260 [?]. The entire audible spectrum contains 30 3rd-octave bands resulting in 10 octave bands. The center frequencies of the bands is specified. Three adjacent 3rd-octave bands can be concentrated into one octave band. It is however not possible to reverse this process as the information of how the acoustic energy is distributed among the three 3rd-octave bands is lost. The exact frequencies can be taken from the appendix.

FFT Apart from using a filter bank it is possible to divide a signal into frequency bands using the discrete fourier transformation, precisely the computationally efficient implementation in form of the fast fourier transformation or *FFT*. The fundamental principle of the fourier analysis is the assumption that every harmonic signal can be deconstructed into basic sinusoidal waves. This is not only true for periodic signals but can even be applied to single pulses. Before the FFT can be applied the continuous acoustic signal has to be

sampled, i.e. converted into time-discrete digital data. A set of 2^n samples is then used as input data for the FFT. The number of input samples used is referred to as block size. The Nyquist criterion and therefore the sampling frequency f_s limits the highest frequency which can be detected using the FFT method. The lowest frequency is determined by the number of samples and therefore the duration of the input data set. Apart from limiting the lowest frequency the number of samples also limits the frequency resolution which can be achieved since the signal is divided into $f_s/2^n$ frequencies. Unlike the bandwidth of typical filter banks (see above) the bandwidth of the FFT bands is constant. Using short blocks of input data increases the time resolution of the output since the length of the block can be considered similar to a time constant for averaging hence short peaks become less distinct for larger block sizes.

No matter how the spectrum of a signal has been analyzed there are two main types of plot, which *figure 2.4* for both 3rd-octave and FFT spectra.

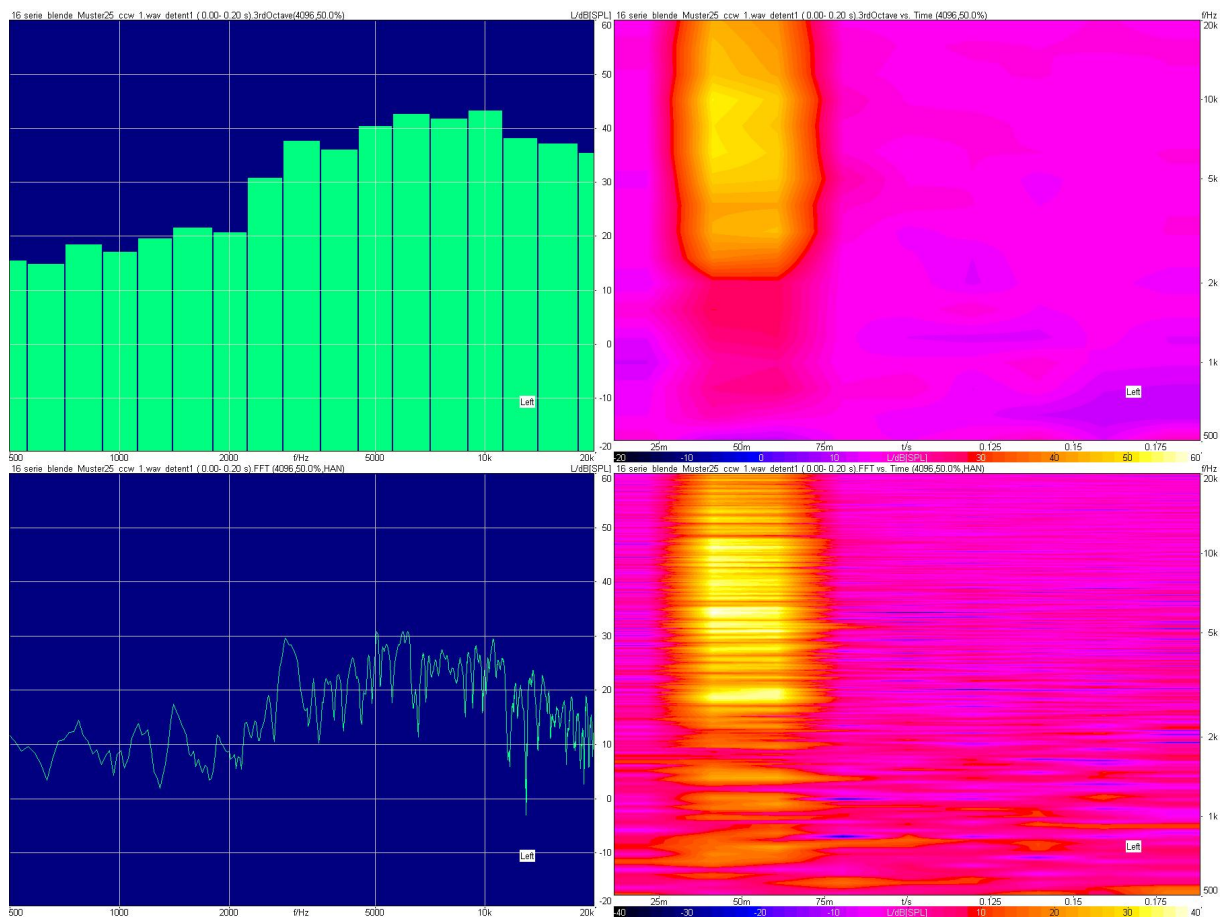


Figure 2.4: The figures on the left show averaged plots, the figures on the right frequency over time plots. The upper plots show 3rd-octave spectra, the lower plots FFT spectra. The input signal is in all cases a click signal of an ecoder.

The figures on the left show averaged values which are displayed as bars. The abscissa denotes the frequency and the ordinate denotes the partial sound pressure level. These figures do not show if there are time dependent fluctuations of the signal's spectrum. The frequency

over time graphs on the right however basically are 3D plots. The abscissa denotes time, the ordinate denotes the frequency and the color represents the partial sound pressure level at any given point on the time and frequency axis.

2.2 Processing by the sense of hearing

This section explains how sound is received by the sense of hearing. Basically, the reception of sound can be divided into two different parts. The first is the preprocessing that happens in the ear itself. The later subsections deal with measurement scales which have been developed in order to represent sensations which are caused by the sense of hearing.

2.2.1 Preprocessing

Outer Ear The propagation of a sound wave is affected by any object which is large compared with its wavelength λ . Head and torso of a subject for instance alter the sound field.

The outer ear canal can be considered as an open pipe. It is about 2 cm long which cor-

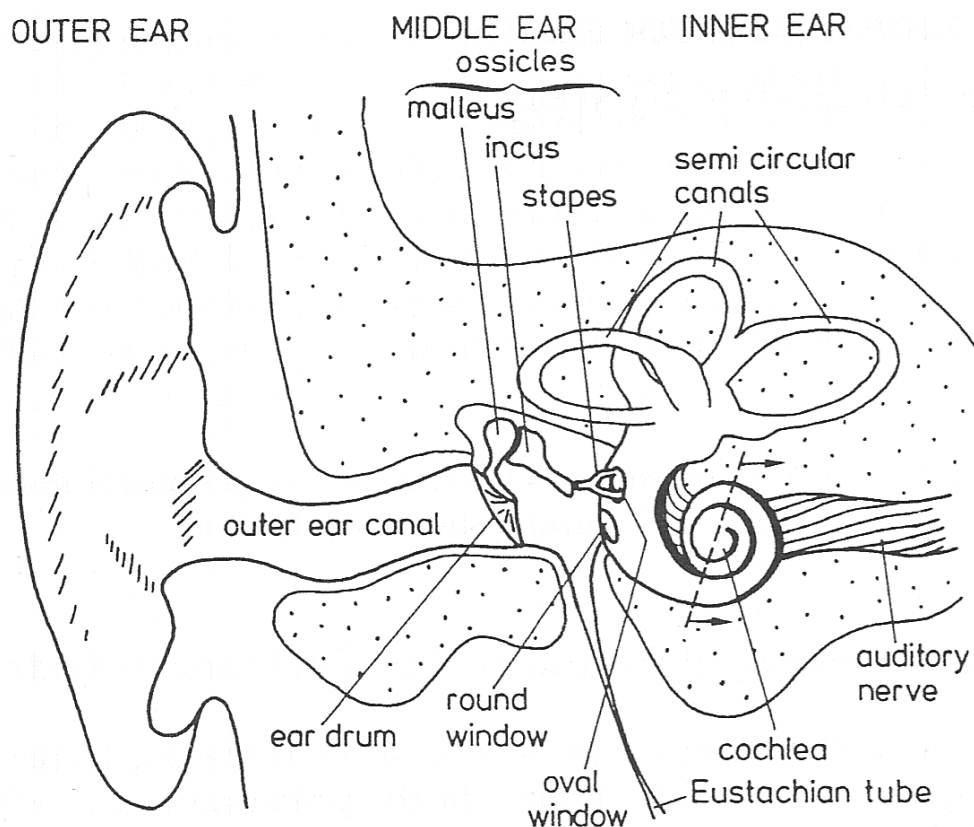


Figure 2.5: Schematic drawing of the outer, middle and inner ear [10]

responds to a quarter of a wavelength for 4 kHz. This explains the high sensitivity of the human ear at this frequency. Apart from the function as resonator and collector of sound

energy the main purpose of the outer ear is to protect the middle ear from damage and to allow the inner ear to be positioned very close to the brain. This results in short nerves hence short travel times.

Middle Ear The outer ear canal is filled with air while the inner ear is filled with a salt-water like fluid. The middle ear's function is the transformation of oscillations with small forces and large displacements into oscillations with relatively large forces and small displacements. Without the middle ear sound waves would be mostly reflected when they hit the inner ear resulting in large energy losses and poor sensitivity. The middle ear acts as a transformer which does impedance matching between inner and outer ear. It works as a system of very small levers. The levers are the three bones malleus, incus and stapes (hammer, anvil and stirrup) which transfer the motions of the ear drum on a membrane called the oval window which is the entrance to the inner ear which is called the cochlea. With a lever ratio of about 2 and an area ratio between ear drum and footplate of about 15 the middle ear achieves almost perfect impedance matching in the frequency range of about 1 kHz. The Eustachian channel is normally closed but opened briefly while swallowing. This brief opening allows pressure equalization between the outside and the otherwise closed chamber in the middle ear. Without this pressure equalization there would be constant tension on the eardrum as soon as the ambient air pressure changes. This constant tension changes the frequency response as well as the sensitivity of the ear. The effect can be observed in airplanes or during diving.

Inner Ear The inner ear is shaped like a snail and consists of three liquid-filled channels or *scalae*. The basilar membrane runs between the *scala media* and the *scala tympani*. Due to the 2.5 turns which are formed by the cochlea the basilar membrane can be about 32 mm long. It is connected to the organ Corti. Its function is the transformation of the vibration of the inner ear into signals that can be processed by the nervous system. This is achieved by sensory cells or hair cells. Sound waves with different frequencies cause different areas of the basilar membrane to vibrate as shown in *figure 2.6*.

High frequencies cause vibration near the oval window whereas low frequencies cause vibration near the helicotrema. It becomes obvious that the inner ear fulfills the task of frequency separation. This function of the inner ear was discovered by Georg von Békésy who was awarded the nobel prize for medicine in 1961. Since vibrations of the basilar membrane show crosstalk to neighbouring hair cells the effect of masking can be explained.

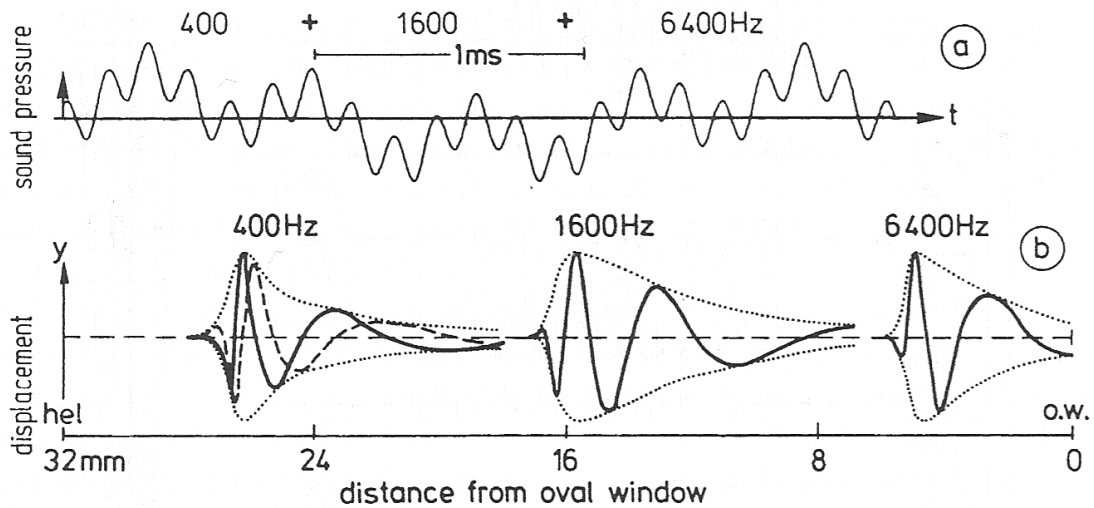


Figure 2.6: Frequency analysis on the basilar membrane. The acoustic signal contains three frequencies. Each frequency excites a certain part of the basilar membrane. However, there is crosstalk to lower and higher frequencies [10].

2.2.2 Perception

Pitch Pitch is a sensation elicited by sound waves. Basically it is related to the frequency of a received sound wave. The pitch of pure tones can be directly measured in listening tests in which a reference tone of a certain frequency is presented. The test person's task is to adjust the frequency of a second tone in a way that it is perceived as for example half the pitch of the reference tone. For low frequencies, for example a reference tone of 440 Hz would lead to 220 Hz which is expected. However, if the reference tone is set to 8 kHz, unexpectedly 1300 Hz are perceived as half pitch.

The behaviour of our perception of pitch can be seen in *figure 2.7*. Pitch was assigned the unit 'mel' which is derived from the word melody. An increase of pitch by the factor two on the mel-scale means a perceived doubling of pitch. The reference frequency is a pure 125 Hz tone which corresponds to 125 mel. The perception of pitch is also subject to effects caused by the level of the signal or noise which is masking the signal. However, these effects are very small and only alter the pitch by 3-5 %, they therefore will be neglected in this work.

While the sensation of pitch can be relatively easily defined for pure tones there, the complexity of the sensation increases with the complexity of the stimulus. Pitch can be detected for all periodic signals as well as noise. However, the pitch strength which indicated how clearly pitch is perceived for a stimulus decreases with increasing bandwidth of noise.

Loudness Just as pitch resembles the way frequencies are perceived the term loudness describes the way, changes in amplitude are perceived. Certainly, stimuli with higher amplitudes are perceived louder than stimuli with lower amplitudes. However, this is only true for stimuli with the same spectral composition and the same length. As stated above the human sense of hearing is not uniformly sensitive throughout the entire frequency range but it is most sensitive between 2 kHz and 5 kHz and becomes less sensitive towards the limits of the listen-

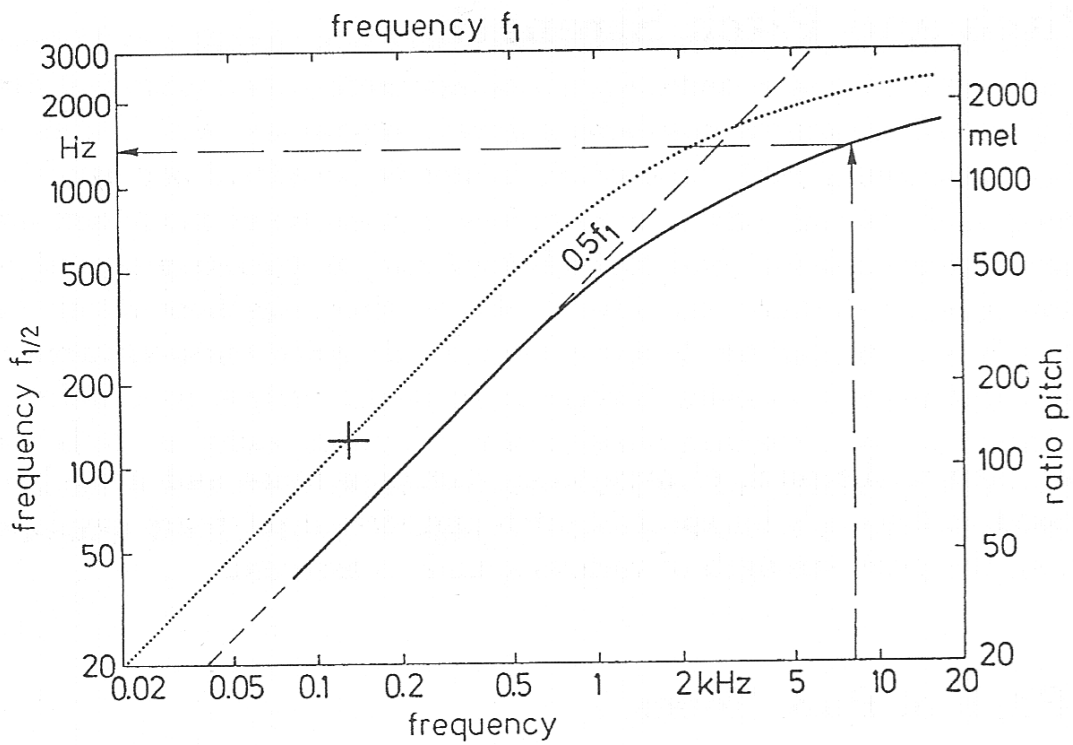


Figure 2.7: Nonlinear perception of pitch. For low frequencies an increase of frequency corresponds to an equivalent increase in perceived pitch. For high frequencies pitch increases at a lower rate [10].

ing plane. Just as pitch, loudness for pure tones can be measured directly by comparisons by test subjects. The reference signal for loudness is a 40 dB 1 kHz tone. This corresponds to a loudness of 1 sone. The sone scale is linear, a perceived doubling in loudness corresponds to a doubling of the sone value. Above a level of 40 dB an increase of sound pressure level by 10 dB means a doubling of loudness. Below 40 dB however, this level difference which is needed to double the perceived loudness becomes smaller until it reaches roughly 2 dB at 10 dB sound pressure level.

Critical Bands As mentioned above the basilar membrane acts as a frequency analyzer. Since hair cells are arranged along the basilar membrane each hair cell basically can only receive a certain frequency. However, since all hair cells are connected to the basilar membrane they are subject to crosstalk when the membrane is oscillating with relatively large displacements. For this reason a loud narrow band noise signal (referred to as the masker) masks a quieter noise signal of a slightly different frequency. The masker temporarily alters the hearing threshold for frequencies close to it's own as shown in *figure 2.8*

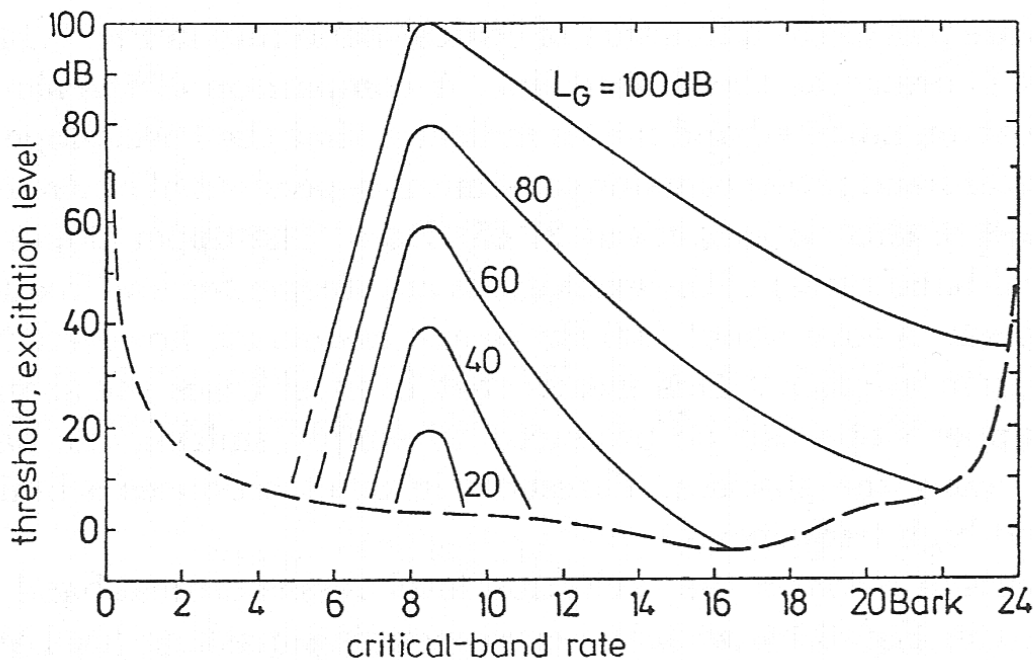


Figure 2.8: Temporary threshold shifts of critical band-wide noise played back at the indicated sound pressure levels. The broken line represents the threshold in quiet [10].

As can be seen the area which is masked increases with the amplitude of the masker. The threshold shifts towards higher frequencies are bigger and extend over a wider frequency range than towards lower frequencies. The basilar membrane can be divided into 24 groups in which frequencies are processed. Just as the mel scale the spectral width of these bands is increasing linearly at low frequencies and exponentially at high frequencies. In psychoacoustic plots sometimes the critical band rate is used instead of the frequency in Hz. The critical band rate ranges from 0 to 24 and has the unit "Bark" (in honor of Barkhausen, a scientist who introduced the "phon" loudness scale which is based on critical bands).

Timbre Apart from loudness and pitch there are further sensory reactions caused by sound which are often generally referred to as timbre. However, a number of psychoacoustic scales has been developed in order to extract specific features from the umbrella term of timbre.

Sharpness is a sensation caused by the spectrum of a signal. Many comparisons indicate that the spectral fine structure of a signal is unimportant. Therefore a white noise signal for example produces the same sharpness as a signal consisting of several spectral lines as long as their power is equally distributed across the listening plane. Furthermore the level dependency of sharpness is very low. An increase of sound pressure level from 30 to 90 dB doubles the sharpness of a signal. The sharpness of a signal is basically calculated by summing up the partial energies for each critical band. However, high frequencies and therefore high critical band rates contribute stronger to the sharpness of a signal than low frequencies. A weighting factor (see *figure 2.9*) has to be applied to the calculation. Sharpness is denoted in the unit "acum" which is the latin expression for sharp. The reference signal is a critical-band wide noise at 1 kHz center frequency and a level of 60 dB.

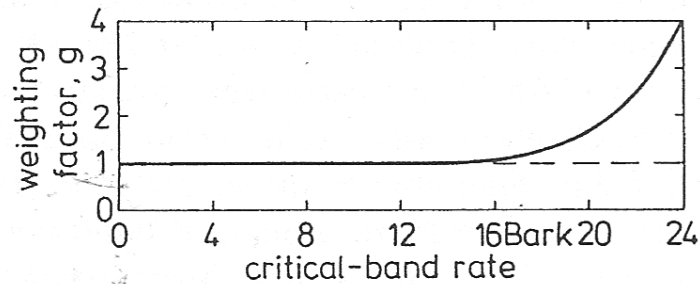


Figure 2.9: Weighting factor for sharpness as a function of critical-band rate [10].

Fluctuation Strength ; describes how modulated signals are perceived. This accords both for frequency modulation of pure tones and amplitude modulation of pure tones and noise. The maximum fluctuation strength is perceived for modulation frequencies of 4 Hz. This corresponds to the temporal envelope of fluent speech: at four syllables per second, which is considered to be the normal speaking rate, the envelope varies at a frequency of 4 Hz. This is a possible explanation for the correlation between speech and the hearing system. The dependance of the sound pressure level is stronger than for the sensation of sharpness. An increase of 40 dB increases the fluctuation strength by a factor of about 2.5. The reference of fluctuation strength is defined as a pure 1 kHz tone played at 60 dB with 100% amplitude modulation at a rate of 4 Hz. This reference signal corresponds to 1 vacil, which is derived from the Latin word vacilare (vacillate in English).

Roughness is a sensation which is also caused by modulation. As shown above fluctuation strength reaches it's maximum around 4 Hz and diminishes at around 30 Hz. For high modulation frequencies the sensation of roughness is elicited. The unit for roughness is called "asper", referring to the Latin term for rough. 1 asper corresponds to a 1 kHz tone that is 100% modulated at a modulation frequency of 70 Hz and played at 60 dB. Roughness occurs

for modulation frequencies of about 15 to 300 Hz. For this reason uniform noise with no additional amplitude modulation causes the perception of roughness.

2.3 Common methods for subjective analysis

A number of methods and procedures are available for subjective analysis of stimuli. They are designed for different psychoacoustical tasks and it takes different amounts of time to arrive at a relevant result.

2.3.1 Paired Comparison

During a paired comparison a number of stimuli are presented arranged in pairs. The subject has to decide which of the two stimuli of the respective pair fulfills the examined criterion better. The method is applicable if effects of different stimulus dimensions (i.e. loudness and pitch) are to be examined. According to the method used by Patsouras [11] the method can also be used to obtain a ranking of sounds of which the features are unclear. By asking the subjects for verbal expressions of way they perceived the sound information for further studies can be obtained. The verbal expressions can for example be used to derive antonym pairs for listening tests using the semantic differential (see below).

2.3.2 Semantic Differential

The semantic differential allows analysis of stimuli on multidimensional scales. The scales typically consist of seven or nine points. Opposing adjectives (i.e. 'bright' and 'dull'), so called antonym pairs, are assigned to the ends of the scale and the subject's task is to mark a point on the scale for each presented stimulus. Basically there is no limitation to the number of adjective pairs which are rated for each stimulus it therefore is possible to examine a number of different stimulus characteristics within a single test. Evaluation of data starts with a correlation analysis of the answers given for different stimulus pairs. This allows reduction of redundant data and thus eliminates unnecessary dimensions of the multidimensional space which is used to finally evaluate the stimuli.

2.3.3 Adjustment

In the method of adjustment the subject has control over the stimulus. If, for example, background noise is presented a possible task might be to adjust the volume of the stimulus until it can be heard. This method is suitable to examine the hearing threshold. Since the presence of noise causes temporary shifts of the auditory threshold (see above) this method allows research of the effects of maskers.

2.3.4 Magnitude Estimation

The method of estimation is useful for direct measurement of the perceived quantity of a signal feature. Subjects have to assign a value to a certain signal feature, for example loudness.

The ratio of loudness can be derived from the ratio of the values. It is useful to present a so called anchor sound. This is a standard which is presented in a pair with the sound that has to be quantified. Sticking to the example of loudness the anchor might be given the value 100. If the subject perceives the sound to be twice as loud as the anchor, the sound would be assigned the value 200. It is also possible to combine the method of adjustment with the method of estimation. In this case the subject would be given an anchor sound and a value for the second sound. The subject's task is it to adjust the second sound in a way that the value fits the sound.

2.3.5 Adaptive Methods

Adaptive methods may be used to obtain information about just noticeable differences (JND) of stimuli. For example the task might be to detect the just noticeable difference in the duration of two stimuli. In this case there would be a reference stimulus of for example 1 kHz and a stimulus which is changed according to the answers of the test subject. After presentation of both stimuli the subject is asked whether the stimulus was perceived higher or lower as the reference. If the stimulus was perceived lower, the frequency would be increased by 100 Hz. After some iterations the stimulus will be perceived as higher. At this point the direction would be reverse and the step size reduces, i.e. the frequency of the stimulus would be decreased by 50 Hz. This repeats until a pre defined final step size is reached. Variations of this method also allow an increase of the step size in case the subject misses a point at which the direction should have been changed.

This methodology has been described by Levitt [12] in detail. Apart from being very useful to obtain JNDs, this method is also suitable to adjust a stimulus to a value which is considered to be good by the subject. This procedure can be applied very well to one dimensional parameters such as the amplitude of a stimulus but with a certain amount of abstraction it is also possible to control more complex parameters in an adaptive way.

Chapter 3

State of Research

This chapter briefly introduces a number of worldwide efforts to understand the perception of transient sounds. This includes approaches how to synthesize these signals in order to obtain controllable and reproducible stimuli for both scientific research and applications. Furthermore projects dealing with the integration of other perception modes, mainly haptics, are presented.

3.1 Analysis and Synthesis of Transient Noise

Common sounds often are no steady-state signals but strongly time dependent. One of the most important examples is speech but also music and numerous sounds generated by machinery are transient. This section focusses on research dealing with transient non-speech signals.

3.1.1 Musical Applications

In Stephen MacAdams essay about the cognitive psychology of music [13] is discussed how pieces of music are perceived. Basically two different structures of perception are outlined in this paper. On the one hand there is the reception and perception of single acoustic events (i.e. each note that is played). On the other hand the entire piece of music i.e. a long series of acoustic events is processed over a time scale which lasts from several minutes to hours. The latter is based on the information which is gained in the auditory system from the reception of single events as well as high level processes in the brain. These processes include the knowledge of the musical form that is being heard and the expectations the listener has to the pieces of music. The perception of those high level structures is therefore affected by cultural background and education as well as the mood of the listener, hence similar to what Donald Norman (see below) refers to as reflective design.

As feedback signals of control elements are not perceived within a large time scale but only as singular events, the knowledge about processing the singular events of music is of interest for this work. Apart from the psychoacoustic values loudness and pitch (see *chapter 2.2.2*), timbre is referred to as feature which is important to classify a singular acoustic event. Timbre is a complex sensory information that enables the listener to distinguish different

instruments which play the same key. Timbre also contains information about the way an instrument is played (for example the stings of a violin which can be either struck or plucked). The essay outlines a synthesis model which uses three dimensions:

- Attack quality - logarithmic rise time of the signal
- Brightness - the spectral composition of the signal
- Global spectral envelope - the time dependent variations of the signal's spectrum

MacAdams cites the work of John M. Grey among others who examined the multidimensional scaling of musical timbres. In his paper on this topic [14] the sounds of different musical instruments are analyzed and subsequently synthesized. This technique of analysis based synthesis allows the usage of stimuli which contain the salient information of the recorded original signals. Recorded signals may differ in a number of different acoustic features. By using synthesized signals as stimuli for listening tests it is possible to normalize all relevant signal features except the one which has to be examined in a listening test. For example it is possible to normalize stimuli in terms of duration and loudness and examine effects of the spectral composition.

Although Grey used the method of analysis and synthesis for harmonic sounds the approach can be transferred to the sounds examined in the work at hand.

3.1.2 Virtual Reality and Games

Robert A. Lutfi uses a similar approach in his works dealing with perception of impact sounds [15]. Like Grey he states that the usage of recorded stimuli or even 'life' stimuli is of limited use in scientific research since the contents of the signal cannot be exactly described. This makes it difficult for other researchers to reproduce the experiment. A variation of the synthesis algorithm for transient sounds proposed by Gaver (see *chapter 6* and [16]) is used in his experiments.

Nevertheless Daniel J. Freed argues in his work on perceived mallet hardness of percussive sounds [17] that the use of recorded signals along with a suitable scale may be superior to the synthesis method since all aspects of the original signal remain. Test results therefore do not depend on the preselection of certain signal features which have been chosen during preparation of a test. Gaver's work also deals with the use of sounds in human machine interfaces. He developed acoustically augmented graphical user interfaces. His research indicates that providing the user with auditory cues in addition to conventional interfaces the use of complex technical equipment such as the control booth of a manufacturing plant can be made easier. He proposed a set of computationally efficient synthesis algorithms to generate so called auditory icons which allow real time synthesis of sounds with variation of a small set of parameters. Apart from the use of this work in user interfaces this work is also of interest for sound design in computer games. The proposed algorithm for transient sounds can be applied to this work.

3.1.3 Sports

G. J. A. Hunter's work on automatic identification, classification and sequence modelling of sound events occurring in tennis matches [18] show a different area in which the analysis of

transient noise is of interest. Impact sounds in sports such as tennis, basketball or golf are emitted by physical events which occur too quickly to be captured by conventional video. As audio signals normally are sampled at a much higher rate they provide data from which this kind of events can be extracted. Hunter uses a hybrid approach to classify events:

At first, acoustic features of single events are extracted. This is done using methods developed for speech recognition such as the calculation of mel cepstral coefficients. Template matching of the cepstral coefficient pattern allows to classify sound events like strokes, echoes and bouncing of the ball. The sequence in which those events occur is used to teach a markov model which supports the classification of sounds. As the markov model is dependent on the context and the sequence in which the sounds occur the same limitations as for perception of music apply. However the ideas to identify single sound events can be used.

3.1.4 Sound Quality

Finally, impulsive sounds are also subject to research in sound quality applications, some recent work also focusses on actuation feedback.

Hard Disk

Modern computer hard disc drives are barely noticeable while idle. The sound of the spinning disc stack is usually masked by other noise sources such as fans. While accessing data however the head of the disc has to be positioned within milliseconds. The highly dynamic movement of the head causes transient noise which is similar to impact noise. By using methods derived from control theory such as input shaping the oscillation and hence the acoustic emission upon repositioning of the heads can be reduced. A. Ali [19] develops a loudness model for the transient sounds caused by hard disc heads for quality control of the drives. Since the development of this model is still in progress there is no further information on this topic available at the moment.

Car Door

R. Liebing uses multidimensional scaling techniques in his work dealing with the noise made by car doors [20]. Liebing uses recordings of actual car doors which are presented using headphones. The test persons rate the presented sounds using 25 semantic differentials. The obtained results are used to derive a calculation model which describes five different sensations which are elicited by the opening and closing sounds of a door. Although the sounds of doors are significantly longer and louder than click sounds of switches parts of the general workflow of this project can be applied to this work.

Touch screens

In recent years touch screen input devices gained importance because of their flexibility as well as cost and space savings as the display area is used as input, too. Lee et al [LINK] examined the impact of multimodal feedback on older adults' performance in a demanding

dual task situation. Merchel and Altinsoy [?] addressed a similar field by examining design and interaction issues for auditory and tactile stimuli in touch displays.

3.2 Applications of acoustics in human machine interfaces, crossmodality studies

This section presents a number of projects which examine phenomena which are related to the subject of this work but cover different fields.

3.2.1 Technical University Dresden - Altinsoy et al.

The work of Ercan Altinsoy covers cross modality studies of tactile and auditory cues. The studies include information on the thresholds for perceiving stimuli in two modes as connected to each other [21]. This research considers both temporal and spatial distance of the stimuli. As for temporal resolution a gap of 25 ms between tactile and auditory stimulus is permitted. The examination of spatial resolution lead to an angle of 6. This means that the sound of the acoustic stimulus has to be placed within 6 of the source of the tactile stimulus. The former has also been examined by Durand R. Begault [4] with similar results. These results have to be considered in the design of the equipment for listening tests. The subjects in crossmodality studies conducted in 2009 by Altinsoy and Merchel [?] performed a dialing task on a touch screen faster and with less errors with auditory feedback which has a temporal structure which differs from what a mechanical control element typically emits. Furthermore they showed in the same paper that tactile feedback has a significant effect on the rate of error in touch screen interactions.

3.2.2 University of Glasgow - Brewster

The work of Brewster et al. [?] considers the crossmodal congruence of stimuli in touch screen applications. Congruence in this respect means that visual, tactile and auditory cues have to be perceived in a way that they match. In a physical switch this can mean that our experiences make us believe that i.e. a big button made from metal has to sound and feel in a certain way while a small plastic button has to feel completely different. If this expectations are not met i.e. the small button sounds "big" and vice versa, we may perceive this as not fitting, hence not congruent. In the mentioned paper Brewster's group used virtual buttons on touch screens in a number of different shapes and assigned different synthetic acoustic stimuli to them with the result that certain stimuli are in fact rated better or worse depending of which button style they are assigned to.

3.2.3 NASA Ames Research Center - Begault

Furthermore Begault examines the applications of acoustics in human machine interfaces for high stress environments such as aircraft flight decks. One of his projects deals with spatial modulation of warning sounds. The results [22] showed that spatial modulation increases the auditory threshold for warning sounds by 7.8 dB. Furthermore Begault's group did research on

adding synthetic acoustic feedback to tools which have to be operated in space by astronauts who - due to the lack of air and noise generated by the life-support equipment in spacesuits - cannot hear the noise of i.e. a cordless screwdriver. Furthermore, he examined the acoustic feedback of tact switches which are operated by users wearing gloves which block the haptic feedback of the switch.

3.2.4 University of Oldenburg - Töpken

The studies of Töpken et al. also deal with cross modality matching of tactile and acoustic stimuli. As mentioned above the works of Begault and Altinsoy state discrimination levels for cross modality matching of these two modes. Beyond the effect of synchrony and spatial effects the results presented in [23] show that also the frequencies and the levels of the stimuli can be used to perceive stimuli obtained through two modes as connected. In his experiments he applied adaptive methods, so the test subjects controlled the tactile stimulus indirectly via their answers until the point of subjective equality (PSE) has been reached. However, the study uses steady-state conditions, therefore it cannot be said if the same effects also occur for transient acoustic stimuli.

3.2.5 Norman

The design critic Donald A. Norman researches how the design of everyday products affects their ease of use. Norman published several books about this topic. In the book "Emotional Design" [8] he introduces a perception model which explains how the design of things is understood. This model consists of three levels of perception:

- Visceral - Things which are understood without any explanation.
- Behavioral - Something has to be explained in order to be understood. After a very short training phase however products with good behavioral design are very easy to use and therefore efficient.
- Reflective - Things require a certain cultural or educational background in order to be understood as intended by the designer. A product which offers its user a high value on the reflective level may be very well accepted even if it is less practical. An extreme example for reflective design is a piece of art without any practical use.

Norman also cites the works of Kurosu and Tractinsky. Both examined the effect of the design of a graphic user interface of an automatic teller machine. Their subjects had to fulfill the same tasks on machines with different interfaces. The time to complete a task was compared to the attractiveness of the respective interfaces. Kurosu's experiment [6] showed that the performance of his subjects increased with the more attractive design. Tractinsky conducted a similar test in another country in order to find out if Kurosu's results are connected to Japanese subjects. His experiments in Israel [7] however show similar results. It can be derived from these observations that it is sensible to design every aspect of a user interface in an attractive way, which includes sound design.

3.2.6 Technical University Munich - Fastl et al

Hugo Fastl [10] is constantly working on the improvement of psychoacoustic scales. His research projects deal with the influence of hearing impairment on the perception of sound [24] as well as the influence of visual stimuli [25] and [26]. During this projects it has been discovered that the color of an item influences the perception of loudness. The sound of a passing train has been presented to test subjects four times accompanied by a still image of a train. The sound was exactly the same every time. The image however has been modified. In one example the Germans ICE trains were presented in their common red and white color scheme, in the other images the train was colored red, green and blue. Results from the listening tests show that the subjects perceived the red train louder than all other trains, the green train was been rated to be the most quiet. To examine if this influence is depending on the content of the sound and the imaged as similar test has been conducted with cars and one with abstract stimuli and color charts. Both led to similar results. So examine the cultural deppency of this effect the studies have been conducted at Japanese universities as well. The results were also comparable to the results gathered in Europe. The influence of color on the perception of noise therefore can be considered to be culturally independent.

3.2.7 Heilbronn University - Reisinger

The work of Jörg Reisinger [3] deals with control elements in cars. Instead of the acoustic properties of control elements tactile cues are examined. The work of Reisinger consists of both the development of the equipment which is suitable to measure the mechanical properties of buttons and switches as well as development of the equipment for subjective analysis and actual subjective studies. The device he proposes for subjective studies (see *figure 3.1*) looks like a usual rotary switch. However, the angle-momentun-characteristic i.e. the haptic feeling of the switch is artificially produced by an electric motor. From the perspective of the test

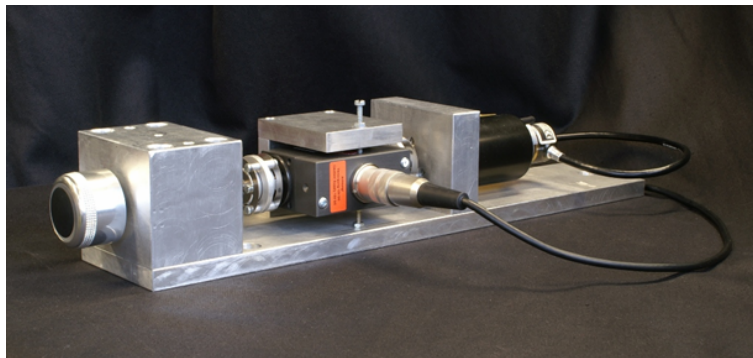


Figure 3.1: Device for haptic evaluation of rotary switches [3]

subjects only the switch can be seen. During tests using this device the subject can not be biased by placement or design of the switch, since all haptic stimuli are presented using the same control element. Reisinger examines the steady-state behaviour of control elements. The transient changes in force or momentum which occur upon actuation are neglected. However, the general methodology of measurement of existing control elements and a device for jury tests which looks like a common switch to present artificial stimuli can also be applied to the acoustics of switches.

3.2.8 AIDE Project

The AIDE project was a collaboration of numerous car manufacturers and universities within Europe which dealt with user usability in next-generation cars. The proposed system (as shown in *figure 3.2*) uses an extensive menu structure. The main display is placed within the driver's line of sight while the system is operated by one or two flat mounted rotary encoders on the steering wheel. The prototype interface in the figure also utilizes a HUD .



Figure 3.2: Prototype of an AIDE interface in a Volvo truck demonstration vehicle [27].

The concept of the system is to reduce the workload of the driver by being adaptive. Data of the current traffic situation is used to run a workload manager which is a software function that prioritizes information and presents it to the driver according to priority of the information and workload of the driver. Driving along an empty motorway at moderate speed might be considered to be a situation with low workload hence, while a winding mountain pass or in dense traffic may be considered to be a high-workload situation which is demanding more awareness. In the latter situation the workload manager filters out data which is considered to have a low priority. Low priority messages include incoming text messages on the mobile phone or a reminder that a light just broke down whereas warnings which require instant reaction such as distance warnings or rapid decrease of tire pressure are still forwarded to the driver. Depending of the urgency of the warning signal the driver can be alerted using acoustic warning signals, warning messages on the display or bright warning lights which are projected on a head up display. The use of these bright lights can be augmented by acoustic warnings and is used in very critical situations, i.e. an imminent collision. The project shows that the multimodal exchange of information is beneficial to the usability of cars.

Chapter 4

Goals and Approach of this Work

4.1 Current Methods of Specifying and Measuring Transient Sounds

As explained in the previous chapter, the specification of transient sounds is important in a number of industrial applications. It is common to simply limit the maximum value of the allowed sound pressure level. In the area of luxury cars desired sounds are usually defined by a small number of experts. Typically the definition is based on the selection of a control element with the desired acoustic feeling from a range of physical prototypes. This prototype is henceforth defined as a benchmark against which samples from the production car are occasionally compared. The comparison is also done by experts.

4.2 Discussion of the Limitations and Shortcomings of the Currently Used Methods

4.2.1 Specification of Maximum Sound Pressure

Depending on the demands a specification based on sound pressure level alone might be sufficient. However, this is only true if it is required that the transient sound is not noticed by a user. The mentioned ongoing research focussing the sound of computer hard disc drives [?] shows that even in this field more advanced methods of measuring and specifying transient signals are sought after. If it is required that a transient sound *can* be heard it becomes obvious that the simple specification of a maximum is insufficient. Musical instruments can be used as obvious examples here:

A pizzicato string sound and the sound of a snare drum can both be recorded and normalized to the same level. Yet the two sounds are clearly not very similar, even if they both are transient. One can argue that the tonal nature of the string and the noiselike nature of the drum is the reason for this. But the same effect occurs for two different kinds of drums and even for the same drum played with different sticks or mallets.

4.2.2 Practical Implementation of Subjective Evaluation

The sounds of control elements are generally noiselike. The sounds depend on the mechanical construction of a certain control element as well as tolerances during the manufacturing process. In the musical instrument analogy the mechanical construction can be regarded as the type of drum and the tolerances can be seen as the sticks. Sound specifications - or sound preferences to be precise are typically stated by a single expert listener or a panel thereof. The panel which defines the desired sound selects a control element which they consider acoustically suitable. Usually, a subjectively very "good" example is picked as a reference. Furthermore one or more examples of "bad" sounds are picked as well. Samples which are taken from the production line for quality control are later on compared by either the same or other experts who judge the acoustic quality.

Expert Listener Panel

While this method usually leads to relatively exact classifications of sounds within a specific panel, the results can vary when the same sounds are compared by other people. This can be an issue in larger companies in which the manufacturing process takes place at other sites than the development. Furthermore, if the decision is only dependent on a single person, the desired sound of the average customer may or may not be in line with this person's taste. Finally, in current development processes prototypes equipped with every possible control element have to be built to enable the expert panel to test the sounds. Building prototypes costs both time and money, prohibiting too many iterations of prototypes to be built. It would therefore be an advantage if

- the specification / selection of a desired sound would be done by as many people as possible
- simple yet effective rules existed to check, if a sound meets the specification without the help of all the people whose opinion the specification is based on
- a solution existed which enables designers to design specifications based on acoustic prototypes

Unwanted Multimodal Evaluation

Apart from the time and cost effect of the evaluation of prototypes there is another constraint when physical prototypes are used. It might be true that a specific control element which sounds best to the expert panel also feels best. But it also might be that an acoustically good sample is rejected because of poor haptic feeling or because of its visual appearance as the impressions of other senses affect the rating. To eliminate those unwanted multimodal interactions it is therefore reasonable to isolate the acoustic impression. Basically, this can easily be achieved by recording the sound of a control element and playing it back. However, in this case the sound is isolated from its context (QUELLE BLAUERT). There are approaches to eliminate these interactions in haptic research tools [?] by building an electronically programmable device. In acoustic research a team at NASA (QUELLE) conducted experiments with subjects which wore thick gloves and were blindfolded in order to impair

there tactile and visual senses. Ideally there should be an acoustic rapid prototyping tool for control elements which is able to do the following:

- Play back either recorded or synthesized acoustic signals of a control element.
- Play back stimuli in an interactive way.
- Play back stimuli in a meaningful context.

4.2.3 Highly Subjective Target Sound Specification

4.3 Proposed Improvements

In order to overcome the outlined shortcomings of the currently common approach to specification and analysis of transient signals the following improvements to methodology for analytics and the approach to specifications are proposed:

4.3.1 Introduction of Spectral and Temporal Properties

As shown above the use of the overall sound pressure level is insufficient to describe the acoustic quality of a signal. The following *figures 4.1, 4.2 and 4.3* will help to understand the possibilities:

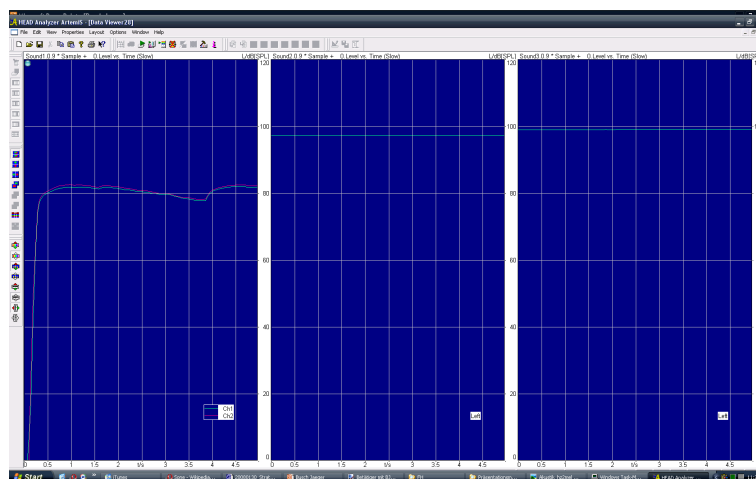


Figure 4.1: Sound pressure vs. time diagrams of three recordings. Based on this information alone one can assume that the three signals are similar

The above figure only shows the sound pressure level of the signal in question. All three signals show no significant fluctuations in their overall sound pressure level over time. However, their differences become obvious in the next plot:

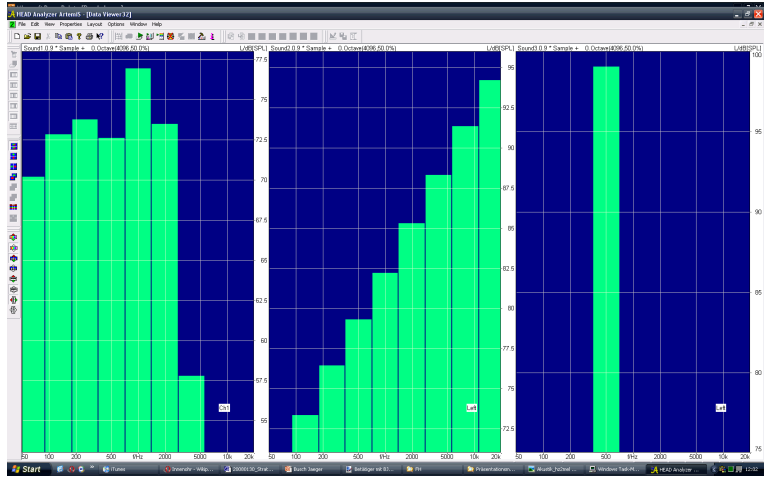


Figure 4.2: Averaged spectra of three sound recordings. It is most obvious that the right graph only shows one significant peak while the other plots seem to be broadband signals

Knowing from what can be taken from both of the above figures one can tell, that the recordings are in fact different but it still is not possible to tell what is contained in the recordings. One can only assume that the two signals on the left are broadband (i.e. noise) and the signal on the right is comparatively narrowband, it could even be a pure tone.

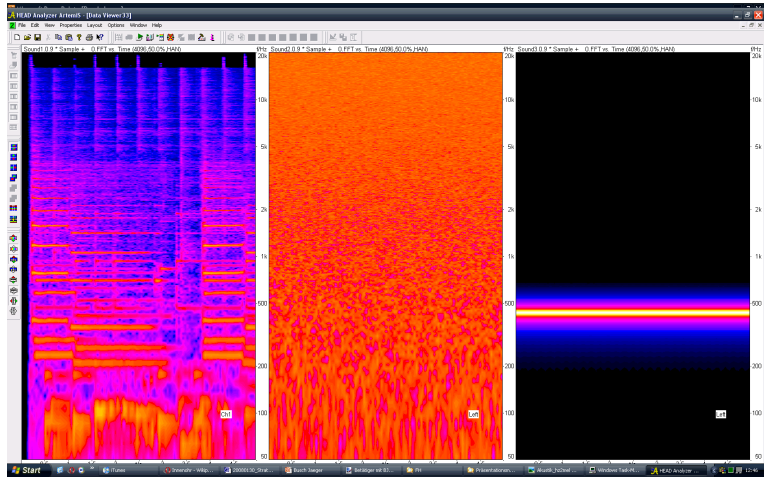


Figure 4.3: Frequency vs. time plots of all three recordings, the temporal structure becomes visible

This last figure now shows the real differences. The picture on the left can be identified as a piece of music, one can see the rhythmic structure and the frequencies (i.e. keys) which are played, the picture in the middle is uniform noise and the picture on the right is really a pure tone.

It basically is possible to record an entire sound, use it as future reference and reject all sounds which do not exactly match this reference without any tolerances.

This would however result in an over-specification of the signal as even sounds which cannot be distinguished by average listeners would be ruled out. Therefore it is proposed to take

into account the temporal structure and the sound pressure level in several frequency bands for a relatively broad classification of sounds. It would be preferable if the used frequency bands are compliant with bands commonly used in acoustic analysis such as octave- or 1/n-octave filters. This would facilitate the implementation of the derived measurement method on industry-standard systems.

4.3.2 Introduction of a Acoustic Rapid Prototyping Methodology for Sound Design

As an improvement to the general approach of selecting and specifying target sounds a toolchain for acoustic rapid prototyping of transient signals is proposed. It consists of four components:

- Recording Equipment
- Synthesis Algorithm
- Interactive Playback Device
- Methodology

The recording equipment is used to acquire the signals emitted by existing electromechanical control elements. The synthesis algorithm enables the user to create desired synthetic feedback signals from scratch. The interactive playback device is used to present the signals within subjective trials. The main feature of the playback device will be that it basically looks and feels like an actual control element but the acoustical character is interchangeable by software. Finally and most importantly a methodological approach to the design of a signal has to be specified.

4.4 Theoretical and Practical Approach

After the challenges of this project have been stated in the previous section the actual approach to the proposed solutions will be outlined. As the human sense of hearing and the emotions evoked by specific noises are highly subjective and vary from person to person it is necessary to involve several subjective experiments in the process. Since even a very short acoustic signal has an almost infinite number of degrees of freedom in terms of temporal structure and frequency, the meaningful parameters have to be identified. This will happen in an iterative series of experiments in which the results and conclusions from one experiment will be used as premise for the subsequent one.

4.4.1 Iterative Examination of Sound Properties in Subjective Trials

Literature (QUELLE) suggests that some parameters (i.e. amplitude) have a larger impact on the subjects' reaction than others and eventually mask the effects of other parameters completely. At the beginning of the project there were no published results about this exact type of sound. Basic research has been done regarding the influence of signal duration on the sensation of perceived loudness. This experiments however were done using signals dissimilar to realistic click sounds. In sound quality analysis it is a common starting point to use existing

real sounds. For this reason the proposed suitable measuring/recording equipment had to be build before the start of the subjective trials. The concept is to use selected recordings in the first trial, extract at least one feature which influences the subjective rating of the sound and normalize this feature in the next iteration.

4.4.2 Isolation of Properties Relevant for Subjective Perception

The isolation of specific relevant signal properties happens after a series of subjective trials is finished and the subjective results of all probands for all sounds in the test have been statistically analyzed. This statistical analysis results in a preference ranking of the used stimuli. The stimuli itself will be acoustically analyzed and ranked in the order of the respective analysis feature (i.e. amplitude, peak/rms ratio, spectral peaks...). Ideally there will be a high correlation of one or more signal parameters to the preference of a sound.

4.4.3 Deviation of a Synthesis Model for Transient Sounds

A synthesizer for the creation of mathematically describable yet realistic sounding stimuli will be developed in parallel and eventually replace the use of recorded signals in later iterations. As the deviation of the synthesis model is an iterative process as well, it will lag behind the isolation of properties by one step. Results gained from a subjective trial which lead to a new identified meaningful signal property will be included into the synthesis model. As soon as the model is sophisticated enough to produce realistic sounding stimuli it can be used for adaptive test procedures in which the subjects have control of the details of the very sound they have to evaluate. This leads to statistical results of the ideal ranges which each of the identified parameters should have.

4.4.4 Deviation of an Analysis Algorithm

After a certain number of iterations the synthesis model will eventually be able to create realistic clicks from a relatively small number of control parameters. Based on the premise that these parameters are sufficient to describe a stimulus the analysis algorithm will be based on the same model. It shall analyze recorded signals in order to resynthesize them using the extracted analysis results fed into the synthesis algorithm.

4.4.5 Experimental Evaluation of Analysis Algorithm in a Subjective Trial

Finally the analysis algorithm has to be evaluated in a subjective trial. The proposed method is to

1. use the synthesizer to adaptively set the desired sound.
2. analyze the subject's desired sound.
3. select sound from the sound database which this subject is likely to like or dislike.
4. present the selected sounds to the subject.
5. compare the predicted acceptance ranking to the actual acceptance.

Chapter 5

Preparation

This chapter briefly outlines supporting work and preparation which had to be accomplished before actually being able to make valid measurements of control elements. It also explains the signal properties of the feedback of a large number of physical control elements, which provide insights on the variation of several parameters when dealing with mass-produced parts.

5.1 Measurement Equipment

The measurement equipment consists of an integrated measurement system for the acoustic feedback of control elements which has been designed and built during the course of this project on the one hand. On the other hand the industry-standard ArtemiS software by head acoustics has been used. This components will be briefly introduced in this section.

5.1.1 Robotic Actuation Unit

Measuring the acoustic feedback of a control element arises challenges which do not occur during the measurement of i.e. and electric motor, ambient noise, a loudspeaker or the vibration of heavy machinery. All of these examples emit their noise constantly without requiring an external mechanical force to drive them. A loudspeaker for example is typically placed in an anechoic chamber, a test signal is fed into the speaker and the measurement microphone picks up the resulting noise emitted by the speaker in a certain distance which is typically 1 m.

A control element on the other hand requires its mechanical operation in order to emit its feedback sound. This in turn requires an external force or momentum to be applied mechanically to the control element. Obviously, in the usual operation this is done by the human user. For measurement purposes however, this solution has certain shortcomings

- undefined angular/linear velocity
- undefined momentum/force
- operator may press a button out of its center accidentally

- operator's hand might be between the device under test and the microphone

While the latter two problems can be solved by sufficiently briefing the operators and careful execution of the measurement, for former two challenges remain as it is impossible for the human hand to turn an encoder for the full 360° at a constant speed and even the constant actuation of a tact switch requires training and concentration, making the use of a human actuator not a very good choice for extensive studies with a high number of devices under test.

Therefore, a robotic actuation unit has been built. The first version which was constructed by Anette Pohl [?] in a diploma thesis project used a stepper motor which directly drives a drill chuck for the actuation of rotary encoders or a linear gearbox for tact switches *figure 5.1*.

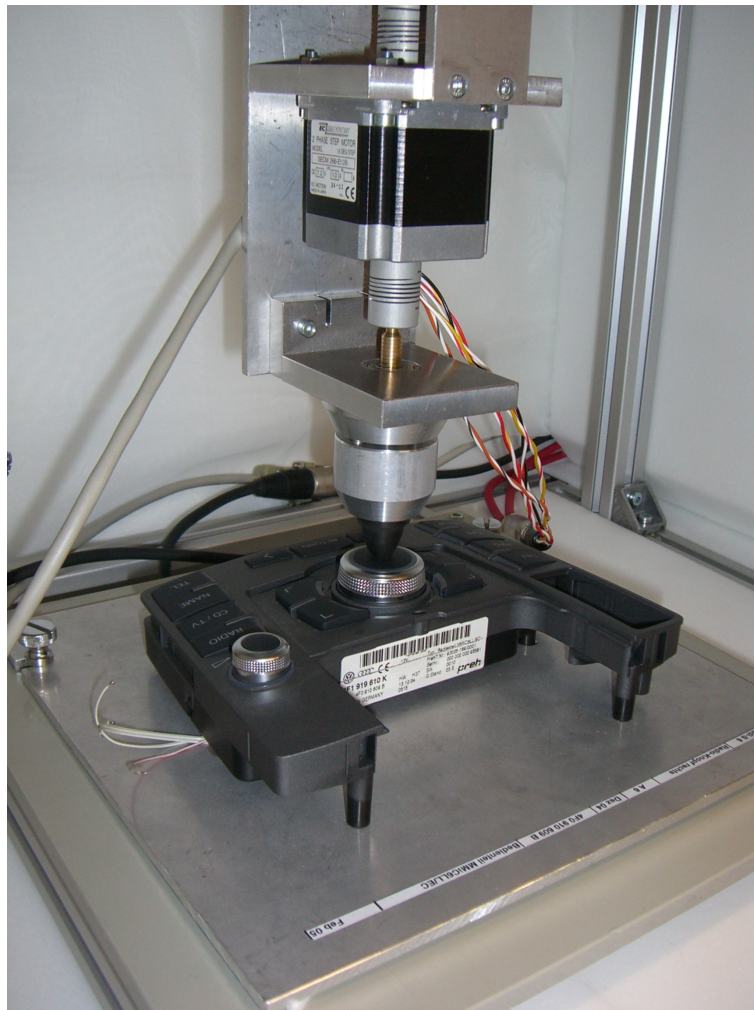


Figure 5.1: First version of the actuation robot in it's anechoic test box

Conventionally, stepper motor are driven with rectangular current signals which allow them to rotate in steps, hence the name. This movement - while easy to predict and control - leads to a very unharmonic movement of the motor and hence to the emission of noise. More sophisticated microstep controllers enable a significantly increased resolution of the steps but

the noise emitted by the motor is still relatively loud in the audible spectrum. Therefore, an analog signal consisting of a sine- and a cosine wave is used to drive the motor. This signal ensures that the motor is driven in infinitesimally small increments and hence quasi-continuous which results in virtually no audible noise.

The revised version of the robot *figure 5.2* uses the same basic layout and motor control principle but the mechanical design has been modified so that entire car components such as climate control panels or radio head-units can be placed in the box.

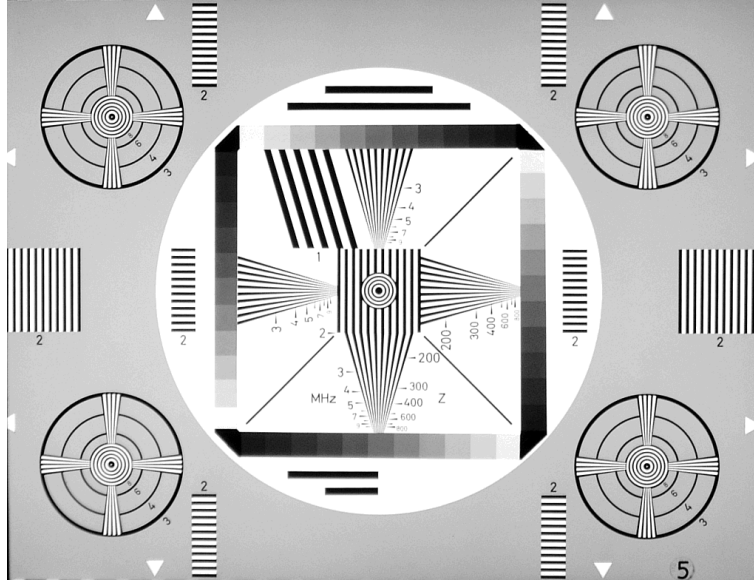


Figure 5.2: Final version of the actuation robot in its anechoic test box

Furthermore the drill chuck has been replaced by a gripper and the linear gearbox has been replaced with a version which features 75 mm travel so that the devices under test can always be placed at the same place under the robot and typical variations in height can be compensated with this gearbox.

5.1.2 Anechoic Test Box

The robotic actuation unit is installed in a specifically designed enclosure which serves as both a case for transportation and storage as well as an acoustically defined space which provides rejection of ambient noise and ensures free-field conditions within the frequency band which is important for the feedback sounds.

The case which is shown in *figure 5.1* measures 932 x 962 x 820 mm. The main idea of the test box is to overcome the need for an anechoic room and enabling the user to do measurements on the desktop. On the outside the box is made from birch plywood which features relatively high internal damping. The wooden panels are covered with sound-absorbing foam (VERWEIS AUF DIE DATENBLAETTER) on the inside. The robotic actuation unit itself sits on rubber dampers for the reduction of structure-borne noise. There are for connectors for sensors (microphones or acceleration sensors) as well as a 68-pin Sub-D connector for general purpose and a 4-pin XLR connector to drive the motor of the actuator. All electrical connections are placed on a terminal which is milled from a single piece of aluminum (see *figure 5.3, 5.4*).

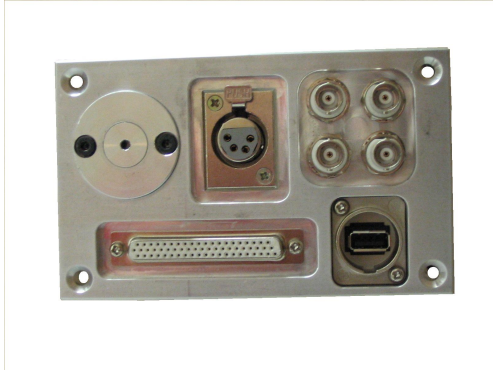


Figure 5.3: Connection Terminal of the Revised Measurement Chamber

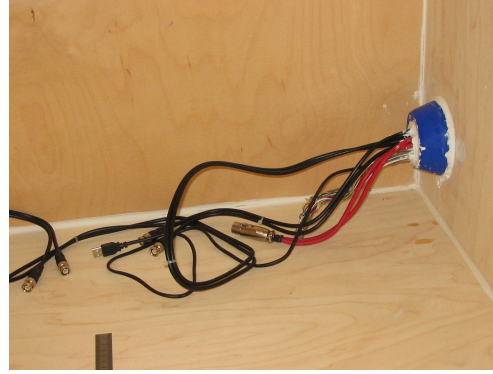


Figure 5.4: Inside of the Connection Terminal

The front panel of the box serves as door and can be opened completely. The hinges of the door can be detached so that the door is removable for work on the inside of the box.

5.1.3 Recording, Signal Conditioning and Data Extraction

The control of the actuation robot as well as the data acquisition, signal conditioning and data extraction are done on a PC connected to the measurement system. The entire software has been designed and implemented in Matlab, however, there is a software version as a result of Reiter's term paper [?] which makes the basic functionality available without requiring matlab.

As mentioned above, the stepper motor is driven by sine/cosine signals which are generated in matlab and subsequently put out using a RME Hammerfall studio soundcard [DATENBLATT] with a RME Multiface DA/AD converter [DATENBLATT]. The same soundcard is also used for data recording. Simultaneous playback and recording in Matlab via a soundcard is conveniently implemented in the pa-wavplay tools by Matt Frear [?].

Due to the acceleration of the motor at the beginning and end of the measurement, the recorded data of the first and last 500 milliseconds are neglected. Since rotary encoders do not have a bedstop, they are usually measured for more than 360° so that still a recording of an entire rotation exists after the beginning and the end of the recording have been discarded. Subsequently, the actual click sounds are extracted from the recording. In tact switches, this is relatively simple as only one click occurs during the recording. The extraction algorithm therefore only detects the peak in the recording and stores 50 ms before and 150 ms after the peak. The measurement microphones also pick up frequencies below 50 Hz by i.e. HVAC systems which cannot be sufficiently rejected by the anechoic case. These frequencies can have relatively large amplitudes and their occurrence can lead to a false detection of the absolute peak of the signal. Therefore, the signal is high-pass-filtered before the detection of the peak. The extraction of the clicks from a rotary encoder basically works in the same way. There are however numerous clicks in one recording which have to be extracted. Typical values are 16 or 20 increments ("*detents*") per rotation but depending on the type of encoder also higher values can be encountered. Aim of the data extraction tool is to extract one recording of every detent of one rotation. For this reason the extraction tool is provided with the information how many detents and hence local maxima have to be detected in the signal. The extractor then detects the first signal peak, extracts it and then detects the next peak in an

interval in which it should occur. This interval can be estimated, as the rotation speed, the sample rate and the number of peaks are known. This procedure is repeated until all peaks have been detected.

The extracted signals are subsequently stored in the standard WAV-format.

5.1.4 Analysis Tools

While custom analysis methods have been developed during the course of this project, the industry-standard ArtemiS software by Head Acoustics has been extensively used as well. This software package can perform a number of standardized analysis methods, including frequency analysis based on FFT or filter implementations, extraction of the psychoacoustic parameters mentioned above as well as analysis methods developed by Head Acoustics which are suitable to detect transient events in noisy backgrounds.

An important feature of the ArtemiS package is that the results are calibrated and therefore valid on an absolute scale. The recording data is stored in the custom HDF format (not to be confused with the *Hierarchical Data Format*). The HDF format is basically comparable with the standard WAV format, which means that the data is stored uncompressed. While the header of a WAV file only stores very little information like sample rate and wordlength, the header of the HDF format contains significantly more information:

- Calibration value for each channel
- Time increment per sample for each channel
- Name for each channel
- Unit for each channel
- Recording date of the file

There is more information stored in the header which can be neglected for the purpose of this study. The head package offers an import tool for WAV files. It is however relatively inconvenient to use if some of the recordings have different calibration values than other or if a larger number of files in a complex file structure has to be imported at once.

For this reason, an export tool for saving the recordings as HDF-file directly from Matlab has been implemented by reverse-engineering the HDF format. A challenge is the recording date and time which is stored in the file. ArtemiS detects if the time in the header differs from the time of the last change which is stored in the filesystem. If this is the case, analysis of the file is rejected.

5.2 Description of the acoustic feedback signal

The acoustic feedback signal of control elements is typically generated when springs inside the switch snap into place. The signal is usually referred to as 'click'. In a typical microswitch (see *figure 5.5*) the spring snaps when enough force for actuation is loaded onto the switch. The spring then connects the contacts of the switch. Sound is radiated upon impact of the spring on the contacts. In the rotary encoder which is used in this work (see *figure 5.6*) the sound is caused by a radial spring which generates both haptic and acoustic feedback. In this section, the properties of the acoustic feedback which occurs in typical rotary control

elements is examined in detail, along with an inspection of the mechanical constriction of one particular type of encoder.

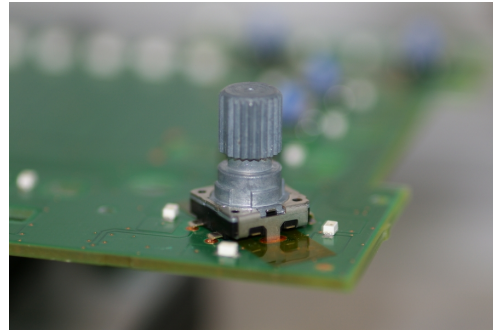
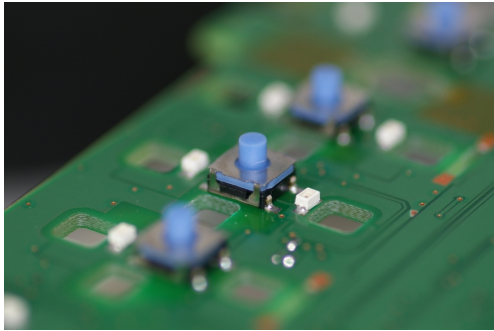


Figure 5.5: Typical micro switch on PCB Figure 5.6: Typical rotary encoder PCB

5.2.1 Amplitude

The feedback sound of a switch can be seen in *figure 5.7*.

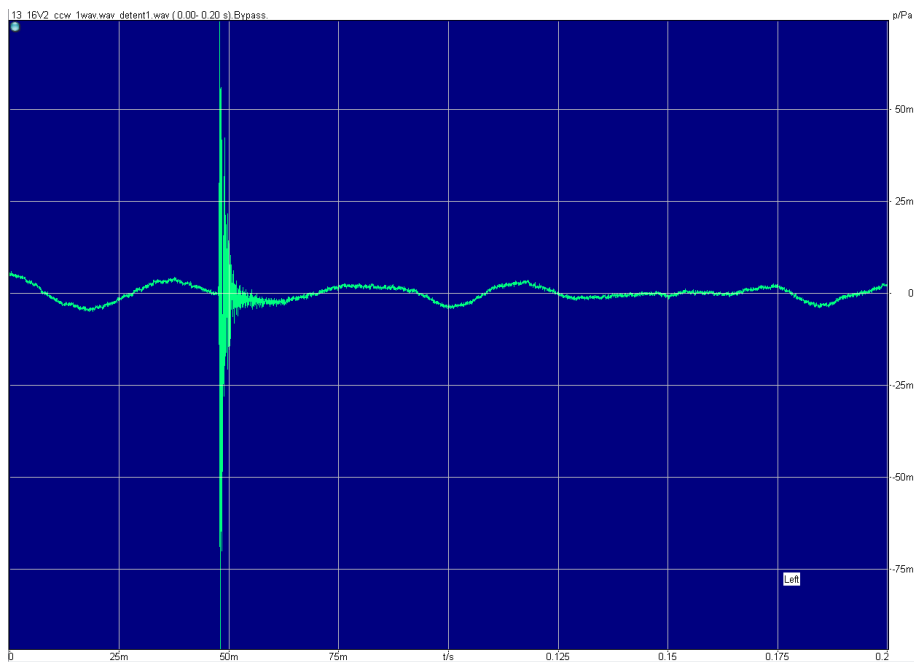


Figure 5.7: Typical acoustic feedback signal ("click") in the time domain

It can be seen that the sound is highly impulsive with only about 1 ms rise and typically 15 ms decay time. Depending on the type of switch these values vary but the length of the signal is in any way shorter than 100 ms which is considered to be the value above which steady-state conditions apply.

5.2.2 Spectral Composition

Due to the impulsiveness of the sounds, the short duration and the small dimensions of typical control elements, the feedback sounds lack frequency components below 500 Hz. A typical frequency over time plot of a rotary encoder can be seen in the following *figure 5.8*.

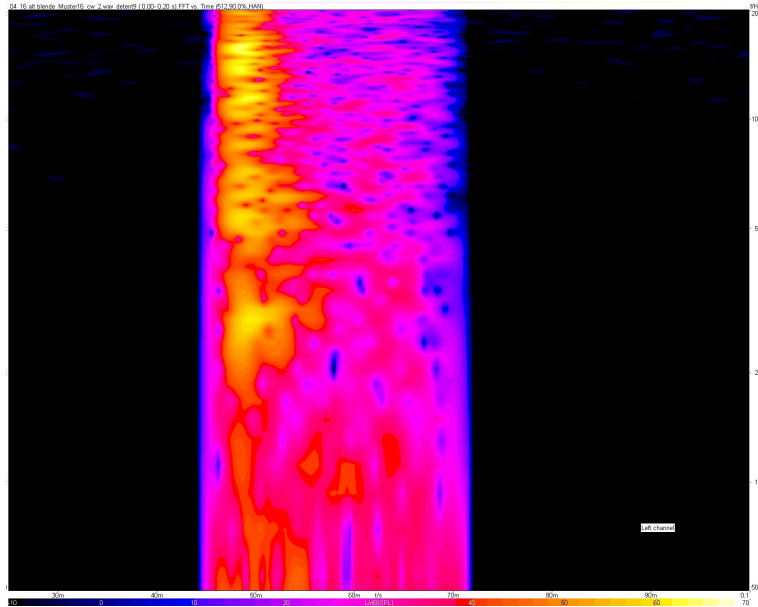


Figure 5.8: Typical acoustic feedback signal ("click") in a spectrogram

In this example the frequency components with the highest intensity are well above 10 kHz, the first local maximum is at roughly 3 kHz. As can be seen in the figure, the maxima are relatively broad on in the frequency domain and that there is no structure observable in the frequency domain. This means, that the feedback signal is not tonal, the spectrogram shows the properties of noise.

5.2.3 Evaluation of Real Control Elements

For the preparation of both the early stages of the subjective trials as well as the evaluation of the proposed analysis method a large number of recorded feedback signals of real control elements was required. For this reason, several hundred rotary encoders have been measured in both clockwise and counterclockwise direction. As each of this encoders features 16 or 20 clicks per rotation this lead to over 30000 recorded single clicks.

This extensive measurement series showed that encoders of the same type show variations in the amplitude of the acoustic feedback exceeding 15 dB as it is shown in *figures 5.9 and 5.10*. The amplitude envelope on the other hand is relatively stable, showing only slight variations within one type of encoder.

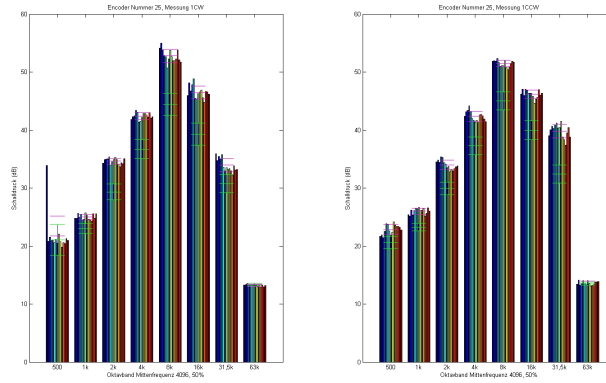


Figure 5.9: Extremely loud sample

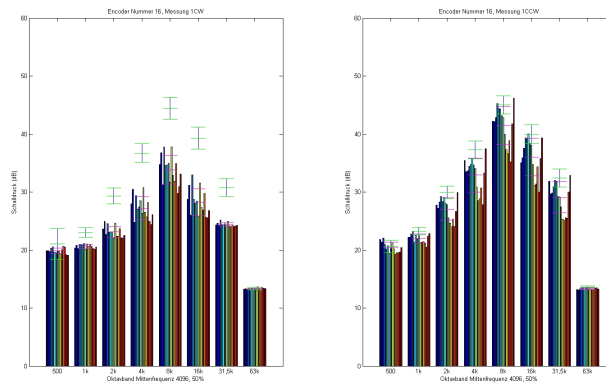


Figure 5.10: Extremely quiet sample

The figures show octave band analysis results for one sample of an encoder. Each octave band bar consists of 16 individual bars which represent the click of one detent. The spectrum is also relatively stable within one type. There are samples which are considered better than others by expert listeners but the overall spectral composition i.e. the relative amplitudes of each frequency band remain similar.

Larger variations of course occur across different types which can happen due to several reasons:

- Different geometry / stiffness of the spring
- Different number of detents and hence changed geometry of the locking disk
- Other Materials used
- Use of viscous dampers (grease)

In the actual use of the control element, there is the even bigger influence of the mounting point of the device. An encoder which is soldered on a very small PCB in a small plastic panel will sound different from an encoder which is screwed to the comparatively large metal case of a i.e. radio head-unit.

5.2.4 High-Speed Video

The following four *figures 5.11, 5.12, 5.13 and 5.14* have been acquired using a high-speed camera system at 1000 frames per second.

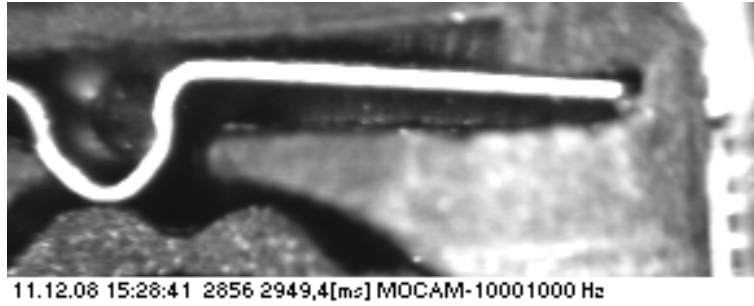


Figure 5.11: The frame shows the state immediately before the movement of the spring

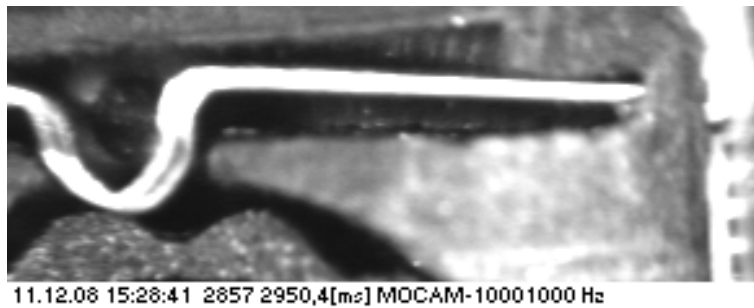


Figure 5.12: The spring is in motion to the right

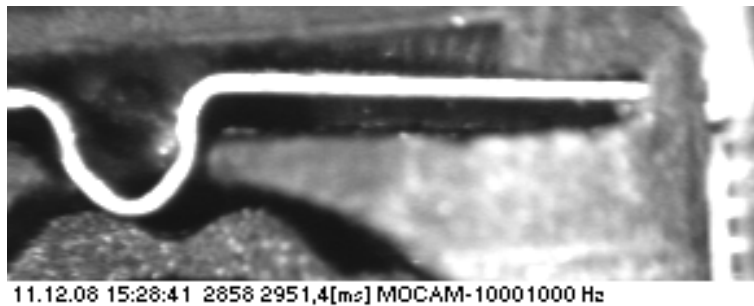


Figure 5.13: The spring has hit its impact

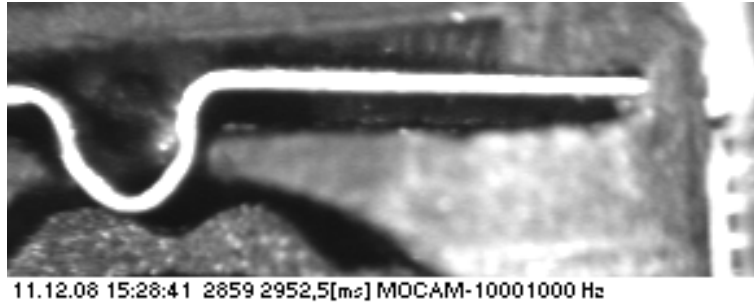


Figure 5.14: No motion of the spring observable 3 ms after the first frame

The camera is pointed at the spring, the first frame shows the spring before the actuation feedback occurs. The second frame shows the motion of the spring to the right and on the third frame the spring has already reached its impact, resulting in an acoustic emission. As can be seen from this pictures, the spring travels less than its thickness to the right and both the radius of the spring and the dimensions of the slot it can move in are subject to manufacturing tolerances. This fact can help to explain the relatively large deviations in amplitude from one encoder to another.

5.3 Interactive Acoustic Simulator

The main tool used in subjective experiments is the so called acoustic simulator. This is a control element with no inherent acoustic feedback. The feedback is generated artificially and played back using loudspeakers. During the course of this project two versions of this simulator have been built. The first version will be briefly outlined whereas the improved version will be explained in more details.

5.3.1 First Version

The first version of the acoustic simulator used a low-friction rotary encoder type Grayhill Series 63K as input device. The encoder impulses are counted using a MiniMEXLE microcontroller board. This microcontroller triggers the playback of acoustic feedback on a Yamaha A5000 (see *figure 5.15*) studio sampler using a Doepfer CTM64 MIDI board [28] [29].



Figure 5.15: The Yamaha A5000 sampler which is used for playback in paired comparisons

The sequence in which the acoustic stimuli are presented to the subjects is also stored on the microcontroller. This experimental setup can be used for paired comparisons of recorded or pre-synthesized stimuli. The subjects respond using a questionnaire. This version of the simulator has been used for the first two experiments but as the test methods became more complex a revised version of the simulator had to be used.

5.3.2 Improved Version

The improved version of the acoustic simulator uses the same silent rotary encoder like the first version. The hardware for test control and response recording however has been changed in several respects in order to conduct more sophisticated tests.

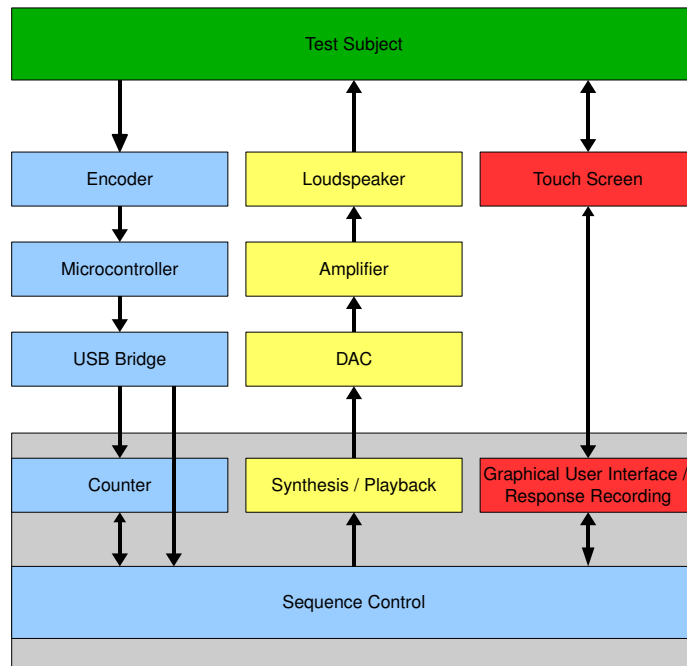


Figure 5.16: Flow chart of the revised simulator

As can be taken from figure 5.16 the test subject turns the encoder which is connected to a custom microcontroller board (see APPENDIX for circuit diagram). The microcontroller counts the increments of the encoder. The encoder used in the simulator provides 256 increments per rotation. Encoders used in user interfaces typically have 16 to 32 detents. In order to trigger audible feedback at realistic angular increments the microcontroller therefore has to convert between the relatively high-resolution signal of the encoder the desired angular spacing between clicks.

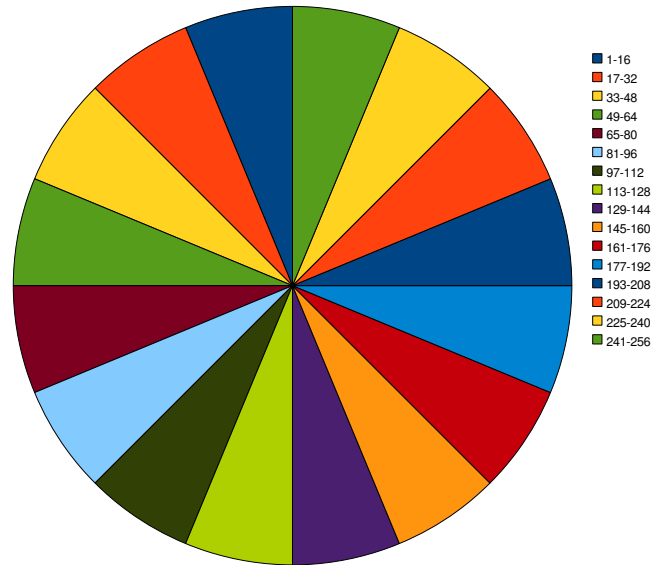


Figure 5.17: This diagram explains how the 256 detents which the hi-resolution encoder offers are used to generate the required amount of trigger signals, 16 in this example

In the case shown in the 5.17 an encoder with 16 detents is simulated. This is the type of encoder mainly used in this research project. Clicks are triggered when the position of the encoder moves from one slice in this pie-chart to the next. The trigger signal consists of one single byte which is sent using an FTDI USB transceiver chip. For clockwise rotation $64+n$ and for counterclockwise $96+n$ is sent whereas n denotes the detent number. This corresponds to the ASCII signs for upper case letters for clockwise and lower case letters for counterclockwise rotation. Each stating with the letter 'a'. The point, that each detent in both directions is an improvement over the first version on the simulator in wich only one trigger signal was sent for every detent.

The trigger signals are received by the simulator software which is implemented in Puredata version 0.40.3.extended, a graphical programming environment for multimedia. The simulator software is used for the control of the test sequence, synthesis and/or playback of the click sounds and graphical user interface. The GUI is used to present tasks to the subject as well as to record the subject's answers and actions. A 10,4" touch screen is used for display and input.

The playback of sound is done using a RME Hammerfall soundcard in combination with a RME Multiface II AD/DA -converter (ANHANG). The sounds are played back at a sampling frequency of 96 kHz and a resolution of 24 bits. The output of the soundcard is connectet to an Hypex UcD180HG amplifier module (ANHANG) using a balanced connection. The amplifier is powered by a Hypex SMPS180 power supply (ANHANG). The transducer is a Scanspeak Revelator D2904/980000 1" Tweeter (ANHANG) which has a useable frequency range of 1-20 kHz.

The transducer is no longer mounted behind the encoder but vertically above it. Since

localization blur of the human sense of hearing is much higher in the vertical axis than in the horizontal axis this simplification can be accepted. The benefit of the placement vertically above the encoder is that the subject's hand no longer shadows the loudspeaker and that resonance issues with the enclosed space between loudspeaker and encoder are avoided. The entire unit consisting of loudspeaker and case containing the encoder and microcontroller hardware is disguised using a black acoustically transparent fabric which is common for the construction of loudspeaker cabinet fronts. This is done to hide the loudspeaker from the subject in order to improve the impression that the sound is emitted from the encoder itself.

Apart from the spatial constraints explained above it is crucial that the sound is perceived within 25 ms after the physical event of turning the encoder (see [21] and [4]). To check if the simulator works within this limit a button on the microcontroller was used to send a trigger signal. The button was connected to channel 1 of a Tektronix TDS 2002 oscilloscope. The second channel was connected to the terminals of the loudspeaker.

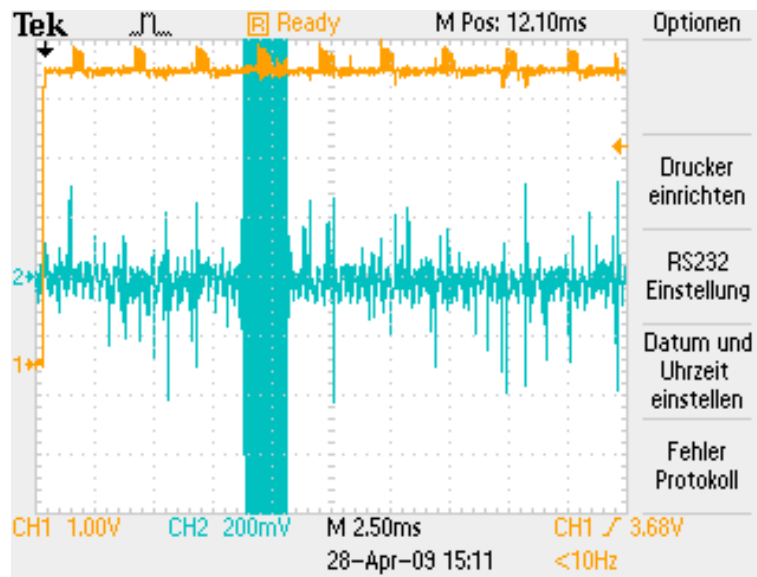


Figure 5.18: Response of the acoustic simulator to a trigger signal

As can be seen in 5.18 the the simulator the loudspeaker receives it's input signal less than 10 ms after the sound has been triggered on the microcontroller board. Since the program to count the increments on the microcontroller is executed every 36 s 5.19 it's impact on the overall response time can be neglected.

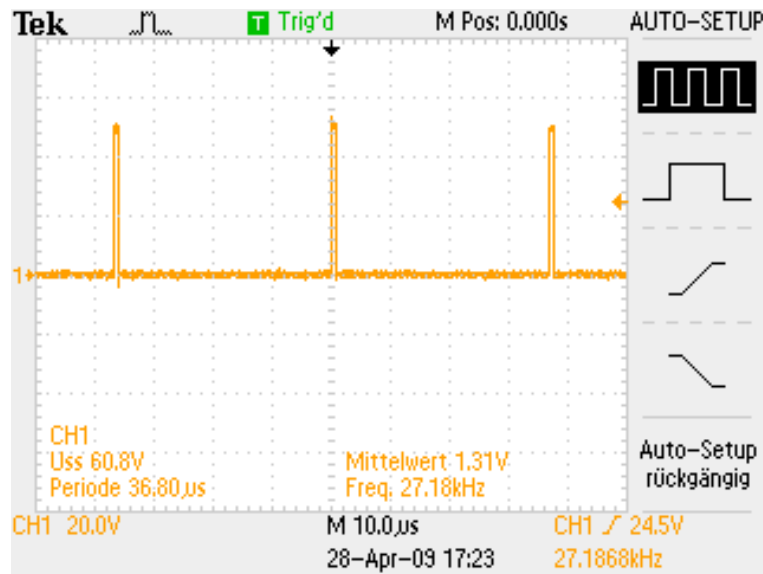


Figure 5.19: Duration of one loop of the microcontroller

To ensure low latency operation of the simulator software Ubuntu Studio has been chosen as operation system for the simulator PC. The system uses the JACK audio environment for communication between the simulator software and the soundcard.

5.4 Listening Room

The simulator itself is placed in a purpose-built listening room, which provides an acoustically controlled environment in which the subjective tests take place.

The subjective evaluation of sound requires a listening environment that allows faithful reproduction of sound. The room basically has to fulfil two criteria.

1. No coloration of sound
2. Rejection of ambient noise

Coloration of sound happens through uncontrolled room resonances and echoes. When a room is excited at its resonance frequency the sound takes longer to die out which means that the reverberation time at this frequency will be longer. This may lead to emphasizing the resonance frequencies of the room. Rejection of ambient noise is obviously needed so that the test subject can concentrate on the presented sounds without distraction by other sonic events. Both criteria are well met by anechoic measurement rooms. However, since anechoic rooms typically have no windows and appear to be threatening to naive subjects a room for jury tests had to be built. The room is designed similar to a recording studio but certain had to be made to suit this purpose. For example, in recording studios there are certain areas in the room with very high sound absorption whereas the opposite of the room is relatively reflective. This is useful to change the acoustic characteristic of the room by altering the position of the artist and / or the microphone. But in the case of a listening room for jury tests a critical design criterion is to have uniform conditions in the room in case more than one person is taking part in a test at the same time.

5.4.1 Layout

The room is located in the basement of the faculty and it's only windows face towards and atrium. Both is helpful to design a quiet room as ambient noise is initially low. Before conversion into the test rooms it used to be a single room. During construction of the test room (see figure 5.20) it has been split up into a control room and the actual listening room.

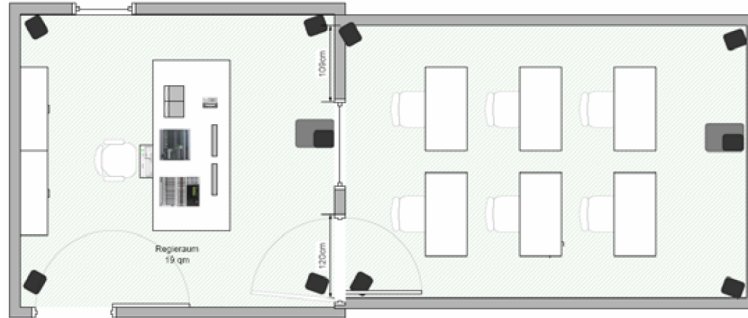


Figure 5.20: Layout of the jury test room. The Control room is on the left, the listening room on the right

The control room accommodates most of the technical equipment such as amplifiers, computer workstations for control, signal generation and documentation as well as the power distribution. It is also the workstation of the researcher who conducts the tests. The control room is 4.5 m wide and 4.5 m long. Both the main entrance and the emergency exit are located in the control room. To access the listening room one has to pass the control room. This ensures that no person can just walk in from the outside and disturb an ongoing test.

Inside the listening room, which ist 4.3 m wide and 6.2 m long, the technical equipment has been kept to a minimum. One of the main design goals was to have no unwanted noise sources in the room. Therefore only the input and output devices of computers (i.e. keyboards, mice, displays and loudspeakers) are located in the listening room.

Both rooms have the same height and both room have an elevated floor and a suspended ceiling for convenient placement of cables.

5.4.2 Simulation and Choice of Materials

Before the construction of the room a digital model has been built. This allowed experimentation with different combinations of reflective and absorbant materials in order to achieve a uniform reverberation time in the listening room. The simulation environment EASE 4.1 offers an extensive database of absorbtion coeffitiens of standard construction materials like concrete or glass. Furthermore an editor allows to add own materials to the database based on datasheets.

The chosen wall absorbers type Illbruck Acoustik "Pyramide 100/50" (refer to *appendix d*) provide an absorbtion coefficient α of over 0.95 for frequencies above 500 Hz. However the predicted reverberation time for frequencies above 1000 Hz was shorter than for frequencies below. Therefore the ceiling elements had to have a decreasing absorbtion coefficient for

increasing frequency. The OWAcoustic Sandila 70 mineral wool ceiling elements fulfil this criterion and were therefore specified.

The process of simulation and the choice of materials has been published at Radioelektronika 2006 [30].

The walls of the control room are partly covered with Primeacoustics absorbers to suppress sound reflections off the walls. All walls are covered with absorbant pyramidic foam.

5.4.3 Acoustical Properties

After construction of the room the T_{60} reverberation time has been measured using MLSSA 10W, an omnidirectional loudspeaker and a 0.5 inch measurement microphone. The loudspeaker is used to excite the room with white noise. As soon as the room is excited the loudspeaker is switched off. The decay time of the sound is then measured. The T_{60} time denotes the time for a 60 dB decay. The results can be seen in the following figure:

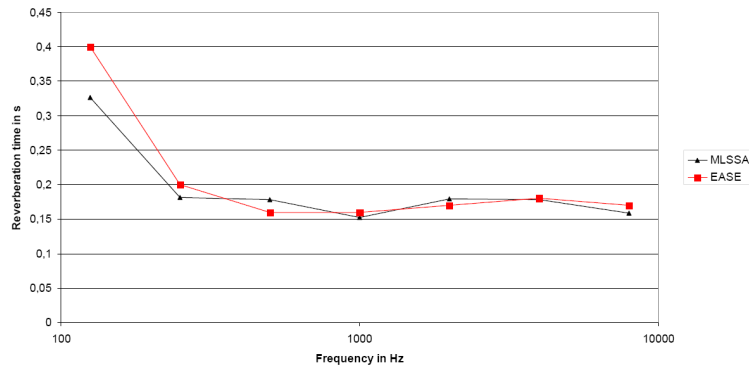


Figure 5.21: Reverberation time of the listening room. The red graph represents the simulation, the black line is the measured result

The predicted results of the simulation correlate with the measured values within 0.03 seconds. The design goal of a uniform reverberation time above 500 Hz has been met. Rejection of ambient noise is above 50 dB.

The acoustical treatment of the control room is less sophisticated since the basic noise level in the control room is already quite high at around 35 dB(A) and its only purpose is to monitor what is going on inside the listening room. The room has been equipped with absorption and scattering elements from Primeacoustics.

Chapter 6

Synthesizer

This chapter describes the basic structure and specific features of a synthesizer for acoustic feedback. As mentioned in the previous chapter, the development of both the synthesizer and the analysis model were iterative processes and ran in parallel over the course of the project. This chapter describe every iteration of the synthesizer in detail but only outlines relevant steps in the development and subsequently thoroughly explains the final version of the synthesizer.

6.1 Basic principle

The basic principle of the synthesizer has been proposed by Gaver [16]. Gaver stated in this paper that impact-like sounds are exponential decays of a sine wave and several hamonics as given in the follwoing formula:

$$G(t) = \sum_n \phi_n e^{-\delta_n t} \cos \omega_n t \quad (6.1)$$

$G(t)$ denotes the resulting waveform over time, consisting of n partials. ϕ_n is the initial amplitude, δ_n the damping constant and ω_n the frequency of partial n .

Gaver cites Freed's [?] work, in which the hardness of a mallet which is used to strike objects was examined regarding the effect on the resulting sound. The ratio of low to high frequency energy in the sounds as well as its change over time served as most powerful predictors for the subjects' hardness judgement.

6.2 Modification of Amplitude Envelope

The described algorithm synthesizes a sound which instantly starts at it's peak amplitude. As the examination of the physical control elements showed, this is not the case in natural feedback sounds. As the work of MacAdams [?] shown, the rise or *attack* phase of the sound of transient musical instruments is critical to their identification.

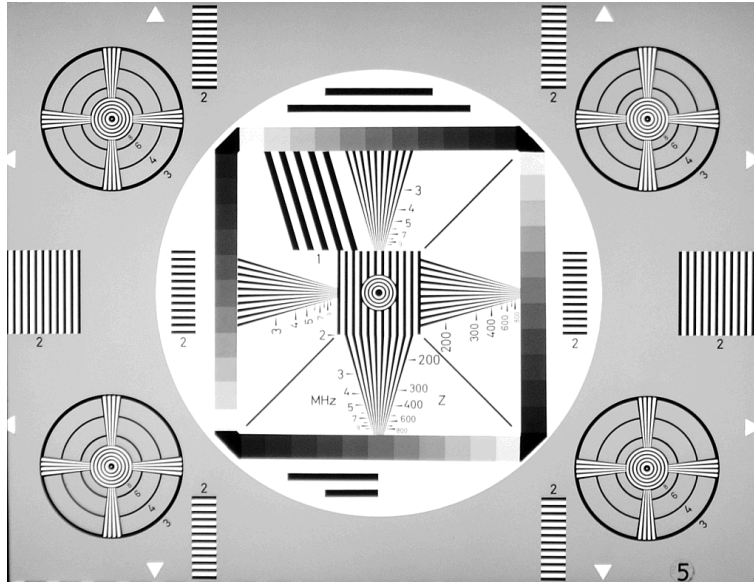


Figure 6.1: Logarithmic Attack

For this reason, Gaver's original algorithm has been enhanced by an attack phase.

6.3 Introduction of Wavetable Oscillator

While Gaver's approach of using harmonic signals works well for impact sounds of i.e. a hard object against a metal bar or a drum head does not produce satisfactory results for control elements. One of the subjective experiments (see *chapter 7.2*) presented in this thesis showed that subjects prefer sounds based on noise over harmonic sounds. For this reason the sine oscillator of Gaver's model was replaced with a wavetable oscillator.

This is basically a concept used in digital sound synthesis and can be used to synthesize sounds based on arbitrary waveforms also shown in *figure 6.2*. The waveform is hereby stored in a table from which values are retrieved and played back.

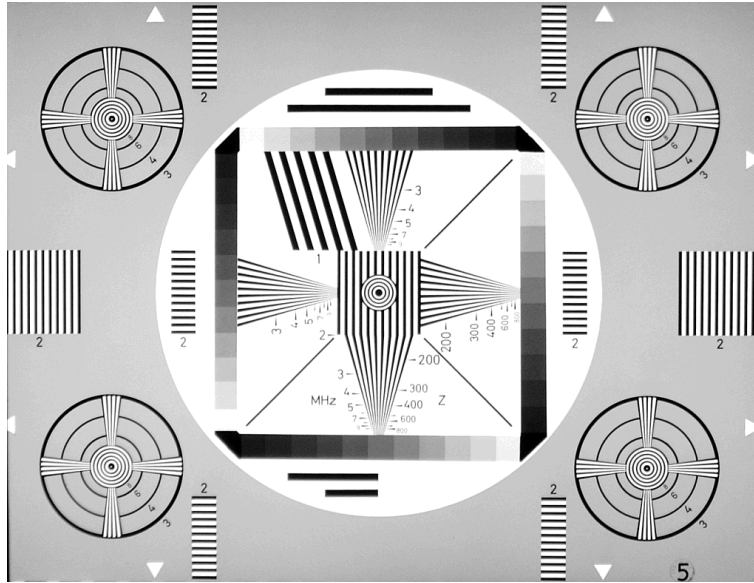


Figure 6.2: Basic principle of Wavetable-lookup synthesis

As seen in the picture, the values are time-discrete and have individual indices. In the example given in the figure in which one entire period of a sine wave is stored in 25 positions or *samples* of the table, the resulting signal would be a sine wave of 1 Hz, if the table would be played back of a rate 25 samples per second. While it is relatively time-consuming to calculate the 25 values of the sine-wave, playback through lookup of a sample requires almost no calculation. Only the sample index has to be incremented, resulting in a ramp which in this example is incremented from 0 to 24 and then starts over at 0 again. The frequency of this so called *phasor* controls the frequency of the playback sound. If, in this example, the phasor would not be incremented by 1 but by 2, the result would be a sine wave of 2 Hz as the table is played back at twice the speed.

As mentioned at the beginning of this section, wavetables may contain arbitrary waveforms and subject's prefer the feedback sounds to be based on noise. This led to a simple implementation with only one oscillator with a uniform noise wavetable. This implementation produces more realistic sounds than the ones created with the original model.

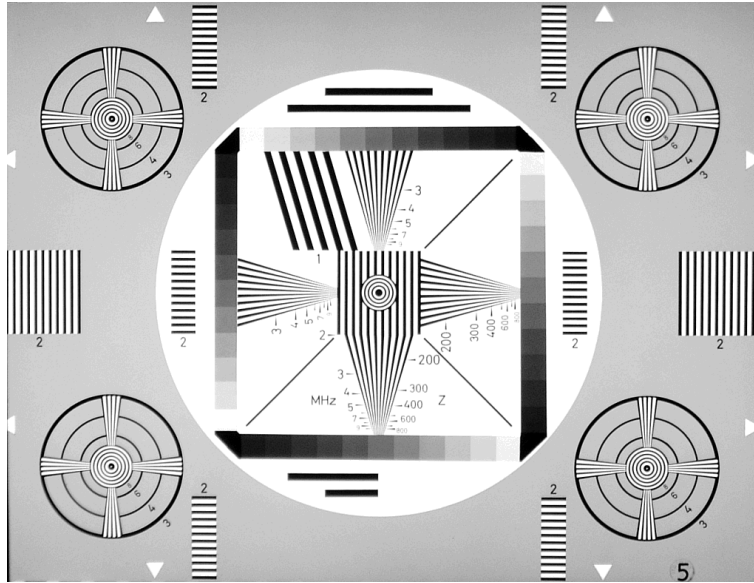


Figure 6.3: Feedback sound synthesized with only one wavetable oscillator

However, the subjects stated in *chapter 7.3* that the spectrum is important to their impression of the sound. In order to gain control over the spectrum the model had to be expanded to multiple wavetable oscillators, each containing band pass filtered noise. To determine the width of the filters it is necessary to understand that due to the limitations of the human sense of hearing only a limited number of frequencies can be perceived simultaneously. Furthermore only one frequency within a given frequency band can be perceived at one time. Blauert (LITERATUR) proposed 23 bands which cover the entire frequency range of human hearing. Especially at higher frequencies (KONKRETISIEREN) these bands are similar to third-octave-bands which are common in acoustic analysis tools (and even consumer electronics and audio software). It is therefore proposed to use the 15 third octave band filters which are defined in the DIN EN 61260 (LITERATUR). The amplitude for each of these wavetable oscillators can be controlled individually. It is now possible to use either one amplitude envelope for a sum of all oscillators or an individual amplitude envelope for each oscillator. The latter provides the possibility to model longer decay times at certain frequency bands which might occur as a result of resonances.

6.4 Control Parameters

While this synthesis model is generally suited to produce mathematically describable synthetic feedback sounds it is relatively difficult to use, due to the number of parameters. Three parameters have to be set for each oscillator, which leads to 45 parameters in total. While this complexity can be handled by an automatic analysis tool which analyses an existing sound to extract its properties for resynthesis, it can not be controlled manually by an untrained test subject. For use in subjective test procedures the number of control parameters therefore has to be reduced significantly. As the results from a subjective test (see *chapter 7.3*) have shown, the actual length of the attack phase, at least within bounds which exist in physical control elements, have very little impact on the acceptance of their sound. Therefore

it has been decided to lock the attack time to a value of 2 ms, which is near the average of a wide number of measured buttons. Furthermore, subjects have frequently stated to like or dislike *dull* or *bright* sounds, rather than *low* or *high* sounds. This means that they do not want to hear very narrow band sounds of a certain pitch but broadband signals with a certain emphasis on low or high frequencies. Rather than allowing the subject to adjust the level for each oscillator it is proposed to simplify the settings by using a model for the sensation of brightness. This can be implemented as a low- or high-pass filter with adjustable slope or by adjusting the levels of each envelope accordingly (see *figures 6.4 and 6.5*).

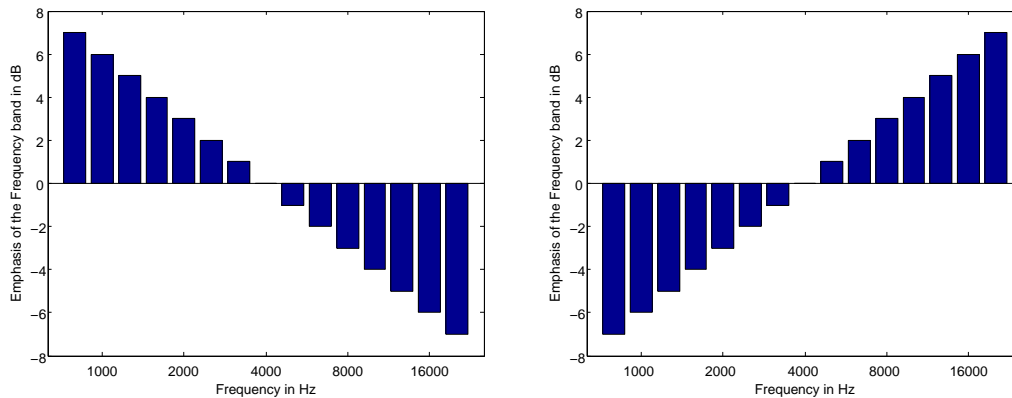


Figure 6.4: Example of a dull spectrum Figure 6.5: Example of a bright spectrum

If the spectrum is adjusted in this way a parameter for the overall amplitude is required, since changes in the spectrum parameter only lead to relative changes of the oscillator levels, not to general amplification or attenuation.

6.5 Synthesis Tool KLKR

The synthesis tool KLKR has been implemented in Matlab. It basically uses the control parameters mentioned above to generate virtual click sounds which sound realistic but are based on a relatively small set of parameters.

6.5.1 Oscillator

The synthesizer uses 15 oscillators which are programmed as fixed-waveform table-lookup oscillators [31]. The reference wavetable for the synthesizer is stored in a wav-file which contains uniform white noise. The program loads a predefined 3rd-octave bandpass filter bank and extracts the 15 3rd-octave signals from this noise signal. The 15 frequency band signals are subsequently stored in 15 columns of an array. Each of these columns represents the wavetable for one oscillator.

6.5.2 build envelope.m

This subroutine builds the amplitude envelope for each oscillator. It uses the attack and decay times as input variables and calculates the logarithmic attack and exponential decay part of the oscillator for each sample of the amplitude envelope. The completed envelope is stored in an array which is returned to the main program.

6.5.3 Phase correction of the oscillator

If the oscillators would be multiplied with their respective envelopes without further processing, errors in the location of the peak of the resulting signal could arise (see figure [?]).

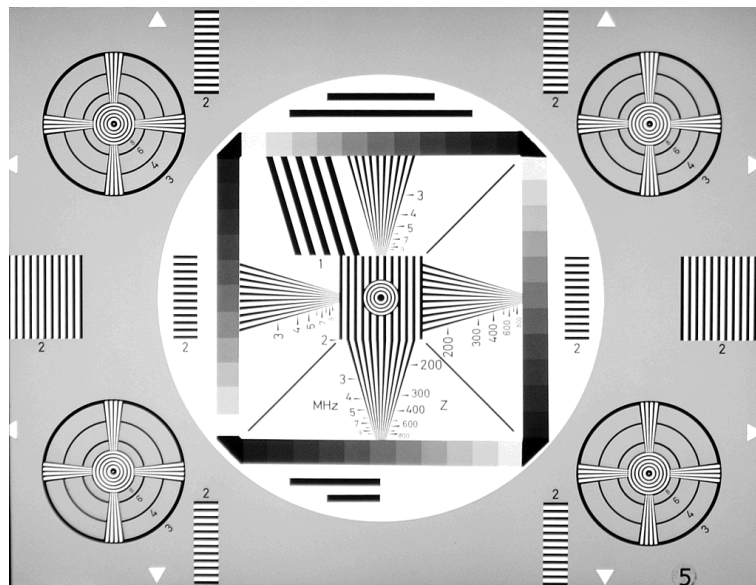


Figure 6.6: BILD DER SITUATION OHNE PHASE CORRECTION, ICH HABS AM 11.2. AUFGEZEICHNET. AM BESTEN EINMAL EIN DREIEBILD UEBEREINANDERN VON EINEM OSC, AN DEM DIE MAXIMA RICHTIG DANEBEN LIEGEN UND DANN EINS MIT KORREKTUR, WO ES PASST.

Since the synthesizer uses 15 independent wavetable oscillators and 15 independent amplitude envelopes with very narrow peaks so that local maxima of the oscillators do not necessarily coincide with the maxima of their respective envelopes. It could even happen that a local maximum of an oscillator coincides with the maximum of the respective envelope. Since all oscillators are summed up in the end this would lead to destructive interference. Since literature and own research suggests that the amplitude of the peak is the most critical factor for perception it is important to avoid this effect. For this reason the contents of each oscillator are delayed or brought forward by the according amount of samples so that their local maxima match the maxima of their amplitude envelopes.

Chapter 7

Subjective Experiments

This chapter explains which subjective experiments have been conducted using the tools and methods explained in the previous chapters.

7.1 Paired Comparisons of Recorded Sounds

In the first series of tests eight recorded clicks of different rotary switches were used. The sounds used in this experiment were chosen by expert listeners from a broad range of data. The data was gathered through a measurement of 150 encoders of five different types. The eight selected sounds (see *figure 7.1*) were chosen because the expert team qualified them as typical for an issue of their respective models.

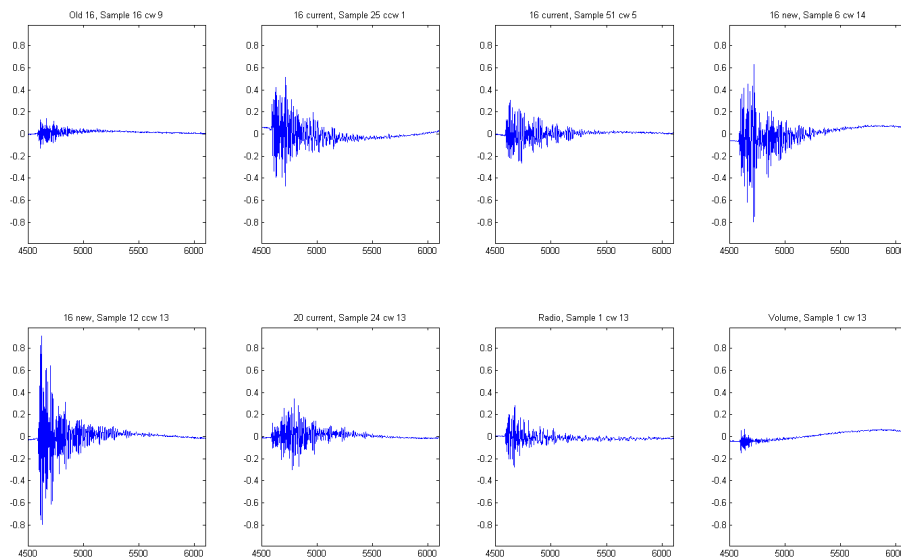


Figure 7.1: All sounds used in the first series of tests

From models which are generally loud extremely loud examples were selected, a model which is regarded as reference the most desirable sound was chosen and so on. All eight sounds were presented on the acoustic simulator. The sounds were grouped in pairs, each sound was compared twice with all other sounds. Two comparisons were done so that each sound appears once as sound A and once as sound B so that an eventual bias of a subject can be compensated. The group of subjects consisted of 30 persons, two of them are experts, 28 are novice and naive to the study. 20 are male, 10 are female. The average age of the subjects was 36. The task of the subject was to select the sound they prefer for every presented pair of sounds. At the end of the study the subjects were given the opportunity to state in their own words what they liked about sounds and what they disliked.

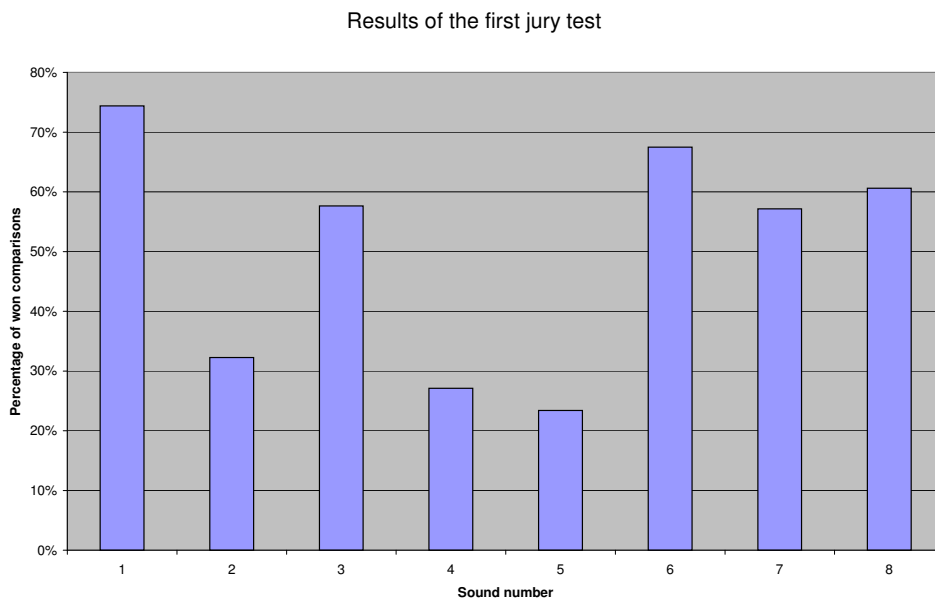


Figure 7.2: Results of the first jury test

It can be seen in *figure 7.2* that the sounds can roughly be separated into two groups. The sounds 2, 4 and 5 were received relatively poor with roughly half the ratings in favor of them compared with the other sounds. According to *figure 7.1* these are the sounds with the highest sounds pressure and also the highest peak loudness. This leads to the conclusion, that the perception of click sounds is correlated to the loudness of sounds. Interestingly, the quietest sound in the range was not rated exceptionally well. The second most quiet sound in the range was perceived best. So obviously there is not only a loudness level that is too high but also a level which is perceived as too low. Furthermore the gaps between the samples within either the "good" or the "bad" group are relatively small and not statistically

meaningful. Analysis of the sounds showed that the same is true for the loudness differences of the samples. As this first series of tests was intended to give the author an overview of the perception of sounds which are meaningful to the panel of experts the sounds have only been chosen on subjective criteria as mentioned above. In order to gain further knowledge on the effects of different acoustic features of the stimuli the sounds for further tests will be chosen based on objective criteria. The list of statements of what the subjects liked or disliked about sounds can roughly be divided into three sections. The first deals with the loudness of the stimuli. Almost half of the subjects agreed that the stimulus is unpleasant when it is too loud. On the other hand five subjects stated being too quiet as bad. Some subjects even stated both. This correlates with the data shown above. The second group of statements deals with the spectrum of the stimuli. About a third of the subjects described a too large part of high frequencies as unpleasant whereas about the same number of subjects described a large part of low frequencies as pleasant. The third part of the statements is rather abstract and cannot be directly connected to acoustic features of the stimuli. Also, most features stated in this category were only stated once. Unpleasant sounds were described as cheap, like plastic, painful, hard, edgy, disturbing and typewriter-like. Pleasant sounds were described as precise, expensive, muted, suitable or like a combination lock of a safe. While the first group of statements supports the data from this experiment the second group dealing with frequency is subject of the second series of tests mentioned in the next section. The third group of words belongs to what Norman [8] [1] describes as reflective level. On the reflective level the appearance of anything is not dependent on a physiological reaction but the cultural background, experiences and education of the test subject. Therefore those answers are highly individual.

7.2 Paired Comparisons of Loudness Normalized Synthetic Sounds

Since the results from the first test series point towards a correlation of amplitude and acceptance the peak loudness of all samples was set to match the peak loudness of the sound which was perceived best in the first series of tests. Goal of the second test was to examine the influence of the spectrum of a click sound on its acceptance. Again the test consisted of eight sounds which had to be compared. One sound was the sound which has been perceived best in the first series of tests. It was used as a reference. The other seven samples have been synthesized in Matlab. The sounds consist of a sample of a generic signal whose amplitude was modulated with an envelope in order to match the amplitude envelope of the recorded click.

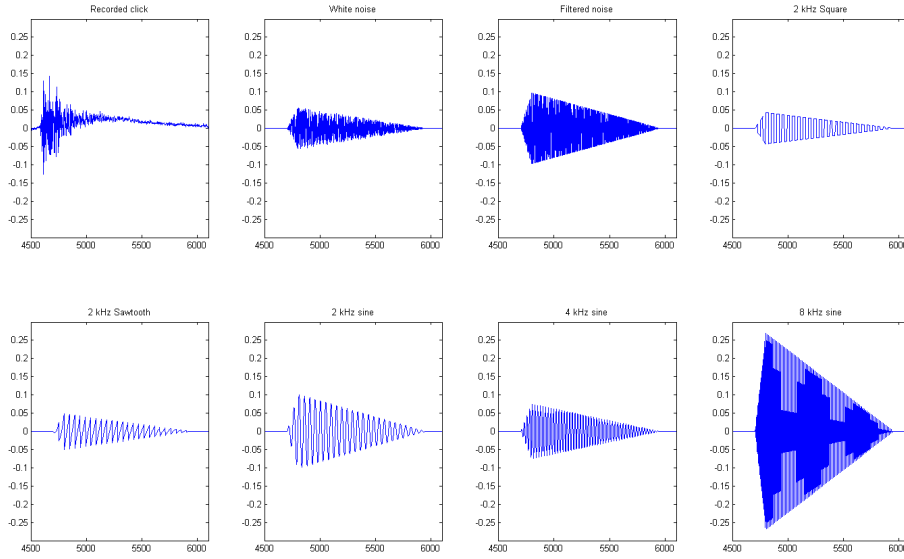


Figure 7.3: All sounds used in the second series of tests

As can be seen in the *figure 7.3* the envelope uses linear attack and release slopes. The envelope is zero at the beginning and the end of the click and reaches it's maximum after a certain number of samples. The magnitude of the maximum as well as the attack and release times are variables. While the release times were fixed throughout the experiment the magnitude was varied to achieve constant loudness throughout the synthesized samples. An overview of the samples can be seen in *figure 7.3*. Note that the peak value of the signal based on the 8 kHz sine is large compared with the other signals whereras the 4 kHz sine samples has the lowest amplitude. The amplitude of the 4 kHz sample is low because of the high sensitivity of the human ears at this frequency (see *figure ??*). The test subjects considered the recorded benchmark sound best. As stated above, in the first series of tests the results cloud be divided into two groups with a relatively wide gap between them. While in the first test series the two blocks had relatively large loudness differences, the two fundamentally different types in this test series were the ones based on periodic signals and the ones based on noise. The recorded reference sound is considered as noise by the author.



Figure 7.4: Results of the second jury test

In fact, the signals based on noise were perceived best by the subjects as can be taken from *figure 7.4*. The recorded click was clearly perceived best. The filtered noise signal has been perceived slightly better than the one based on white noise. The three stimuli based on signals with a 2 kHz base frequency were perceived worse than the ones based on sine waves with the relatively high frequencies of 4 kHz and 8 kHz. Interestingly the 4 kHz stimulus was strongly rejected by the subjects. As explained in *chapter 2.1.1* the human sense of hearing is most sensitive at this frequency. This fact however is taken into account when the loudness of a sound is calculated and since the loudness of the stimuli was normalized prior to the test this is no explanation.

7.3 Rating of Synthesized Sounds

The first two experiments indicated that the acceptance of a sound is amplitude dependant and that broadband signals are perceived better than tonal stimuli. However, the experiments did not provide information on the effect of the temporal structure of the stimuli. For the third experient a more complex synthesis method was used. It is based on the algorithm proposed by Gaver [?] but has been extended by an attack phase. Furthermore the sine osciallators have been replaced with wavetable oscillators which use narrowband noise as wavetable. 20 Oscillators have been used to create the stimuli for this experiment

The stimuli for this experiment were synthesized using analysis results from an extensive database of over 30000 recorded click sounds. The sounds in the database were analyzed regarding their attack and decay times as well as their spectra. The 5 and 95 percentile values for attack and decay times were subsequently used for synthesis. The sounds from the database were recorded from three different subtypes of Greyhill 62AG encoders. For each of the subtypes the spectra have been averaged and. Since Gaver's algorithm suggests the use of a damping parameter to shorten the decay times of high frequency components two arbitrary values have been taken as damping coefficients. In total this results in 24 possible variations of the parameters attack, decay, damping and spectrum.

Since a full paired comparison of this many stimuli would result in 576 pairs a different approach has been chosen. The subjects were ask to rate each sound separately on a nine-point scale ranging from extremely good to extremely bad. Each sound was presented to the subject three times. As the subjects usually not made use of the full nine points of the scale the results were normalized at the end of a test, setting the lowest rating to 0 and the highest rating to 1. The impact of the decay times to the rating of the subjects is pointed out in the following *figure 7.5*:

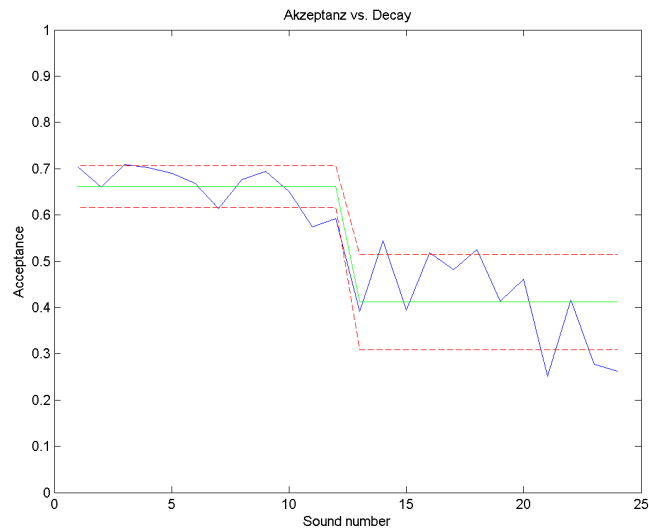


Figure 7.5: Sound Numbers 1-12 use short decay times, sounds 13-24 use the long decay setting. The green line denotes the average for each sound group, the red dashed line denote the standard deviation.

As can be seen, the average acceptance of 0.66 was significantly higher for the sounds which featured the short decay time with an average acceptance of 0.41. This was universally true for all spectra and attack time settings. A reason why this can be observed can be found in basic psychoacoustical effects: the perceived loudness of a stimulus is reduced if it's duration is decreased because of the working principle of the sense of hearing as an integrator.

As the next *figure 7.6* points out, the influence of the spectrum is far smaller:

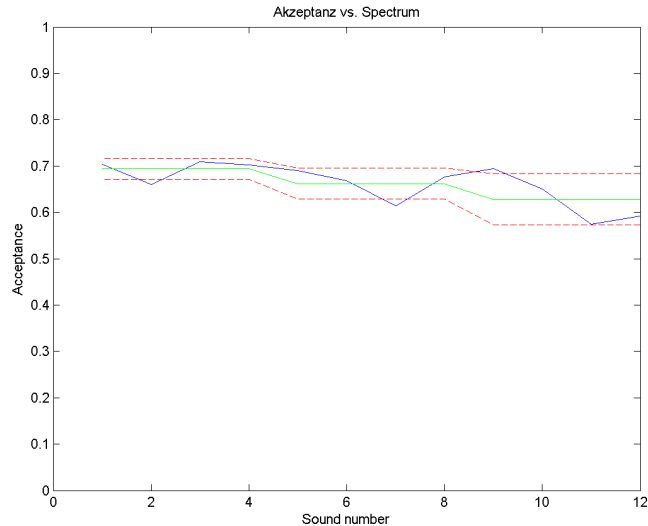


Figure 7.6: The sounds 1-4, 5-8 and 9-12 are based on the same spectra. Especially sounds based on the third spectrum shown relatively low acceptance combined with a relatively high standard deviation.

The figure shows that the sounds based on two of the three spectra received slightly higher ratings than the third group, furthermore the standard deviation in this third group was relatively high.

7.4 Search for JNDs and Reaction Time Measurements

This experiment does not contribute directly to the goal of finding a method for the prediction of the acceptance of feedback. It does however identify, how large the differences of a feedback click have to be at least so that the subject notices the difference. In the actual application of control elements this a relevant question since certain functions such as the temperature setting of the air condition exist twice in most cars (for driver and passenger). The two buttons are normally very close to each other and therefor are expected to feel the same. Furthermore, due to reasons of the mechanical construction of specific switches the sound can change when a switch is turned clockwise or counterclockwise.

Furthermore, this experiment contains a second part in which the subjects had to adjust the sound in a way that they like it. This individually "good" sound was used in a subsequent trial in wick the subjects had so move the control element to defined positions as fast and as accurate as possible. The performace with the good feedback sound was compared with the performace with lacking feedback and the performance with a sound which has been rated very badly in the second experiment.

7.4.1 Design of Experiment

Part 1

The first part of the experiment is used for the determination of just noticeable differences. The control software of the simulator was set up in a way that different stimuli for clockwise and counterclockwise operation of the knob were created. Three signal parameters were examined in the following order:

1. Overall Amplitude
2. Decay Time
3. Complexity of the Spectrum

The approach for each parameter was the same (see *figure 7.7*). In the beginning, the sound in both directions was the same.

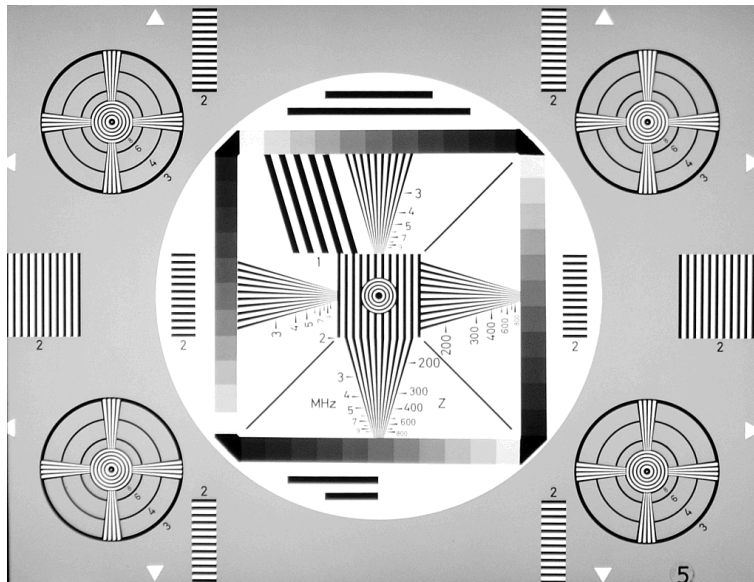


Figure 7.7: Flowchart for the search for just noticeable differences

The subject had to turn the knob in both directions and was asked, if he perceived the sound to be the same. If the answer was yes, the sound was marginally changed. This procedure was repeated until the proband replied that the sounds sounded different. This result was stored, the sounds were reset to the default settings and the next parameter was changed and checked. For evaluation of the amplitude steps the signal was changed by XXX db, for the decay time by XXX percent. For the test of signal complexity, the quietest of the 15 oscillators was muted after the first iteration and the next-most quiet was muted every iteration thereafter.

After all three parameters were checked, the test procedure was interrupted and the subjects were briefed for the next part of the experiment.

Part 2

The this part uses a multiple 2-AFC procedure according to [?] to allow the subjects the manipulation of the signal parameters. The interface is shown in *figure ??*.

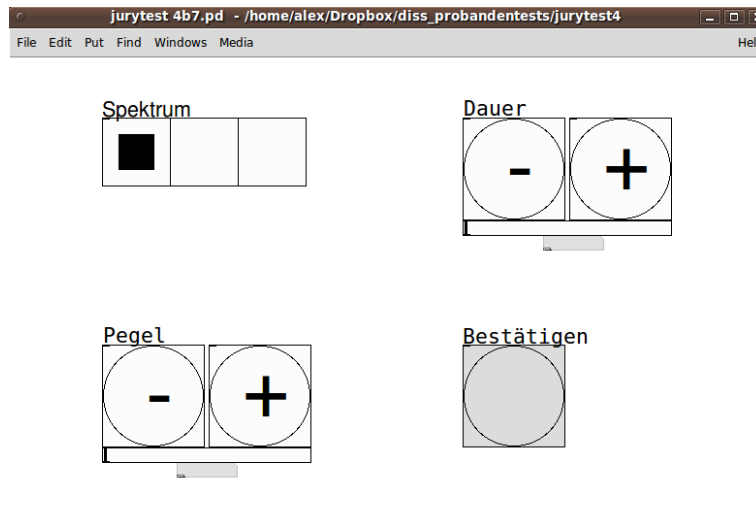


Figure 7.8: User interface for the configuration of an individually ideal sound

As can be seen from the figure. There are four signal parameters which can be adjusted in total:

1. Overall Amplitude
2. Decay Time
3. Attack Time
4. Spectrum

While the first three parameters are self-explanatory, the spectrum parameter differs. At the current stage of the subjective trials, the spectral model for quick access to all amplitudes of the 15 oscillators has not yet been implemented. Moreover, the three spectra from which the probands could choose were based on the spectra of recordings of physical control elements. The procedure for the three other parameters was a common one up one down procedure which reduced the step size by which the respective parameter was adjusted with each reversal of direction (see *figure 7.9*).

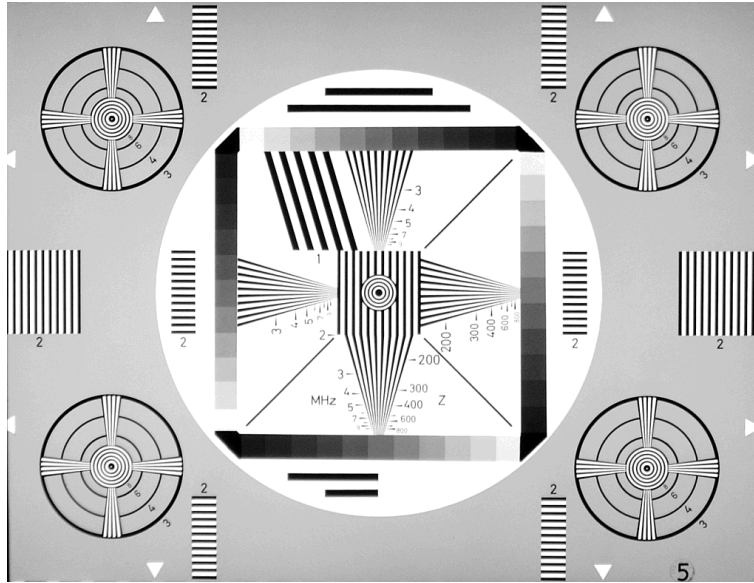


Figure 7.9: Generalized explanation of the one up one down test procedure

Since all parameters affect each other, the subjects were given the opportunity to edit the settings for each parameter even after the final step size has been reached. This means, that the setup of the sound was not automatically finished when the final step sizes for each value had been reached but that the subject had to confirm manually that he is now satisfied with the current setting. This individually desirable sound was stored and subsequently used in the third part.

Part 3

Finally, the last part of the experiment challenges the subjects to move as quickly and accurately as possible to given positions with the control element of the simulator. The subject is only shown the interface shown in *figure 7.10*.

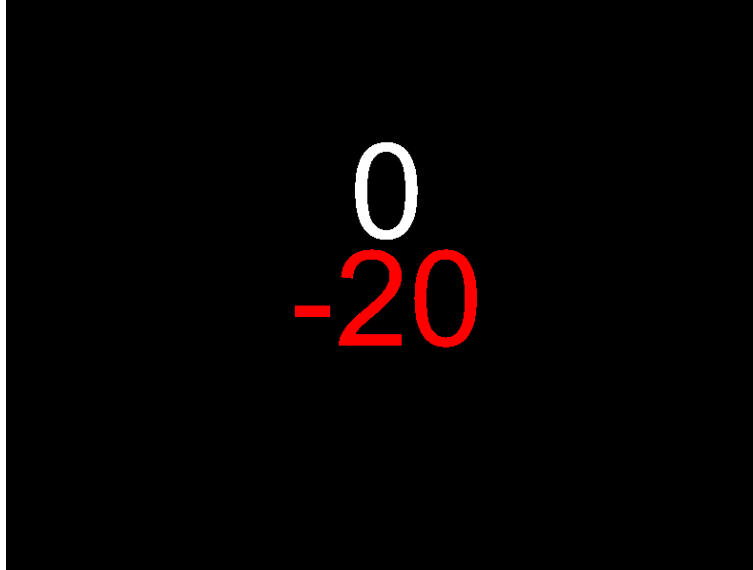


Figure 7.10: Example for task in the third test stage

As can be seen, the interface consists only of two bold numbers. The lower number shows the value to which the upper number has to be adjusted. Positive values require clockwise rotation while negative values require counterclockwise rotation. As soon as the target value is reached the subjects are required to confirm by pressing the space bar of the control computer and the next target value is displayed.

There are ten different target values (1; 3; 5; 10; 20, each positive and negative). The values have been chosen since they represent realistic values for navigation through actual menus in user interfaces. Low values are likely in case one of a few options from a list has to be selected (i.e. in the main menu), while the higher values are likely in navigation through extensive music collections or to enter addresses in the navigation system. Each task is presented three times with different acoustic stimuli:

- The desired acoustic feedback from part 2
- No acoustic feedback at all
- A 4 kHz sine burst at 80 dB(A)

During the preparation of the experiment a table with all 100 tasks has been created. The first 10 tasks are primers which are used to familiarize the subject with the tasks. They comprise all possible values from -20 to 20 and present all three stimuli in an alternating order. After the primers there are 3x3x10 timed tasks, all possible values in all possible combinations with the three stimuli in three repetitions. The timed tasks have been arranged in groups of ten tasks in a row to which the same stimulus is assigned. Within one group, all ten possible target values occur. However, the order in which the target values appear has been randomized so that the subjects cannot anticipate the next target and thus corrupt the results. The stimulus changes in the next block of ten and repeats after three blocks. Since the probands are likely to become faster in fulfilling the tasks during the course of the trial, the assignment of the stimuli to the respective task is rotated after every subject, an excerpt of the table containing the tasks can be seen in *table 7.1*

Task Nr.	Target Value	Stimulus for Subject n	Stimulus for Subject n+1	Stimulus for Subject n+2	Stimulus for Subject n+3
11	-5	1	2	3	1
12	10	1	2	3	1
13	-20	1	2	3	1
...
21	-1	2	3	1	2
...
31	3	3	1	2	3
...
41	1	1	2	3	1
...
100	20	3	1	2	3

Table 7.1: Example of the assignment of stimuli to tasks, note that the assigned stimulus repeats after three subjects, changes every ten tasks and repeats after 30 tasks.

By using this method of assigning stimuli in a rotary pattern to randomized tasks the training effect can be ruled out as long as the number of probands is a multiple of three.

7.4.2 Results

Just Noticeable Differences

Individually Configured Sounds

Evaluation of the data from the second part of the experiment, in which the subject's task was to design their individual feedback signal provides information on how an auditory feedback of a rotary encoder *should* sound like. *Figure 7.11* the individual values for decay-time and peak amplitude which were set by the subjects.

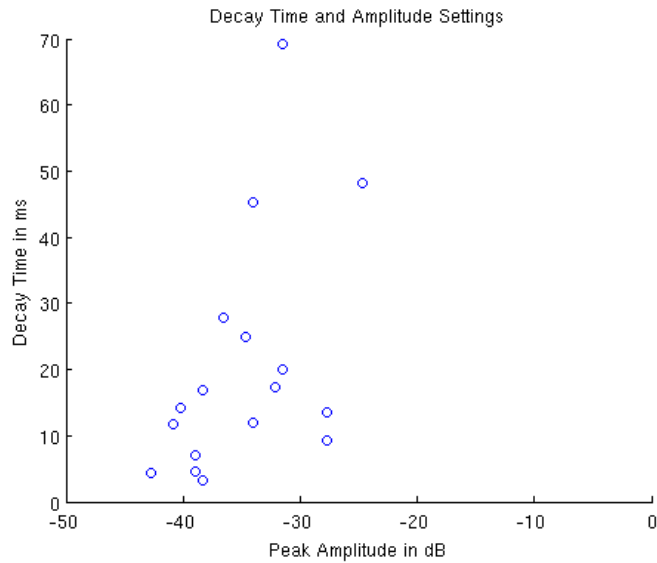


Figure 7.11: The majority of subject performed fastest with the self-defined auditory feedback signal

One can see that the majority of the subjects set the sound to decay within 30 milliseconds and set the peak amplitude within a 10 dB range. The scale on which the amplitude values are plotted uses the full scale of the sound card as a reference are not calibrated regarding actual sound pressure level.

While 10 dB means a perceived doubling of the perceived loudness of the stimulus it is in a similar magnitude than normal tolerances which occur in typical electromechanical encoders due to tolerances in the manufacturing process. Furthermore, two thirds of the subjects preferred one of the three available spectra, as *figure 7.12* illustrates.

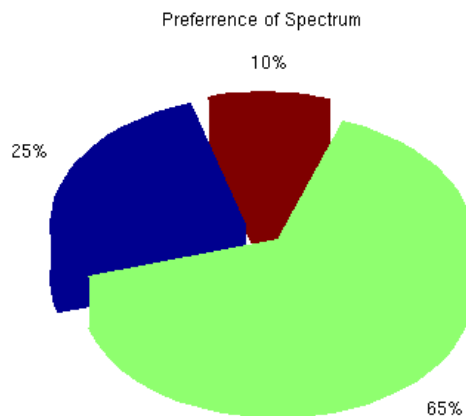


Figure 7.12: The majority of subjects preferred one spectrum

One spectrum on the other hand was almost completely rejected by the subject, although all three spectra were based on a existing electromechanical encoders. This shows that even if the impact of the spectrum on the actual acceptance is small compared to the impact of the amplitude as was pointed out in experiment three, there are certain predominant preferences among all subject regarding the spectrum, at least among this relatively small selections of pre-defined spectra.

Impact of the Acoustic Feedback Quality

Main focus of the evaluation of this study was the time the subjects needed to adjust the control element to the target positions. The following *figure 7.13*. shows the distrubution of probands who performed fastest with a given stimulus.

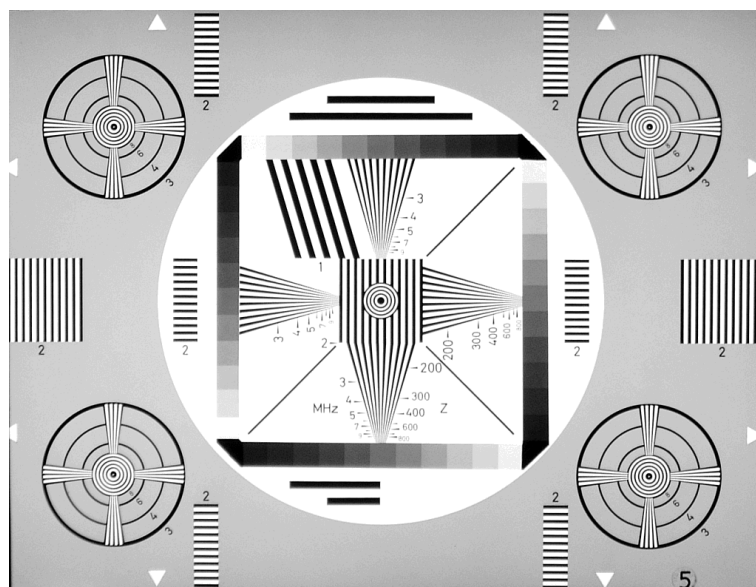


Figure 7.13: The majority of subject performed fastest with the self-defined auditory feedback signal

It can be seen from the figure that the majority of 13 out of 21 subjects were the quickest with their individually set-up sound while only 3 subjects were quickest without any feedback. On average, subjects were 74 ms quicker when feedback was present. Interestingly, the subjects are on average 33 ms quicker with their individual feedback signal than with the very loud but annoying sound. This proves that it is not only interesting from a marketing point of view to provide pleasant feedback but from a usability perspective and ultimately from a safety perspective as well. While the relative gain in task speed seems neglectable as the average task duration was about 2500 ms, the absolute gain is still relevant in an automotive environment since it means that the driver can focus on the road again a sooner. At typical motorway speeds this time amounts to approximately half a car-length. Regarding the accuracy of the subjects it is also the self-defined stimulus that leads to the best performance as can be taken from *figure 7.14*.

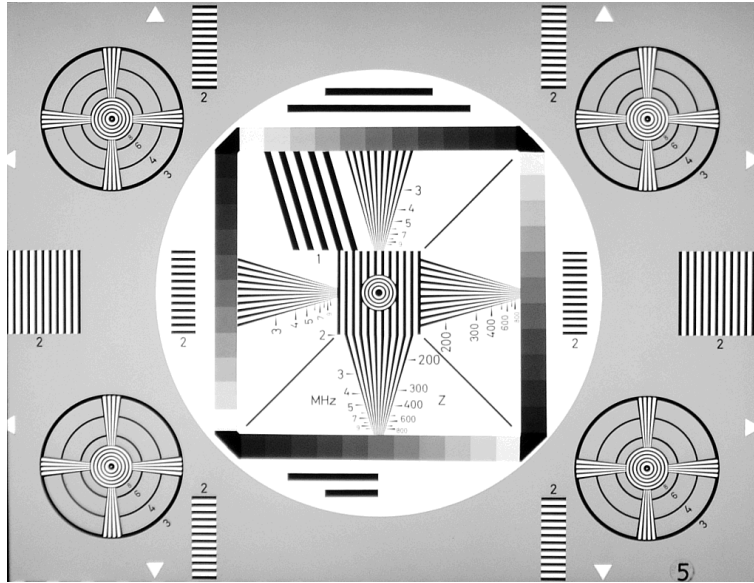


Figure 7.14: The majority of subject performed most accurate with the self-defined auditory feedback signal

To analyze the accuracy, the absolute dialed in value is subtracted from the absolute target value, i.e. it would have been an error of 1 if the proband dialed in 6 instead of 5 but also if the proband dialed in -6 instead of 5. This has been done since some subjects occasionally turned the knob in the opposite direction than required by the sign. Since this experiental setup did not use any meaningful graphic interface which had to be operated through the input device but only used abstract numbers it is valid to use the absolute values. The average errors are shown in *figure 7.15*

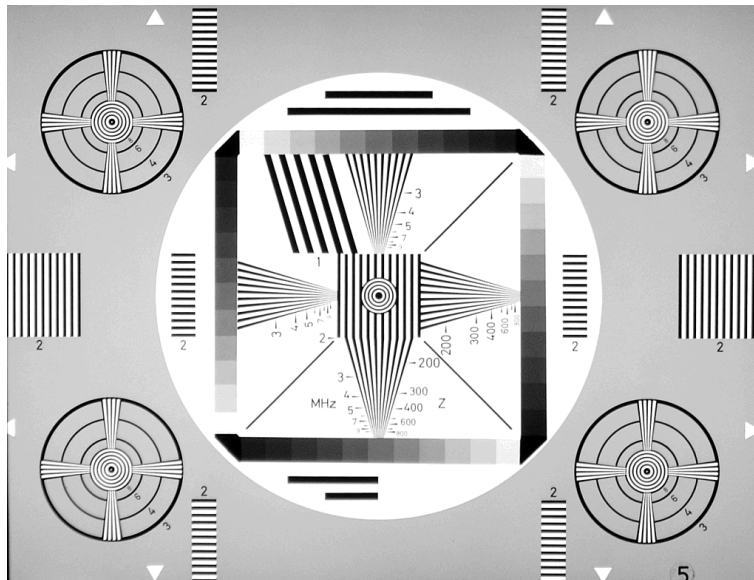


Figure 7.15: Average deviation of the dialed-in value from the target value depending on the used adutory feedback

As can be seen in the figure, the subjects typically performed best using the self defined feedback while the clearly audible but unpleasant feedback increased the rate of error to almost twice the result using the individually optimized stimulus.

In total, the effects auditory feedback on the performace of subjects selecting values using a rotarty encoder as an input device can be summed up in the following *talbe 7.2*

Acoustic Feedback	Mean Error in Detents	Mean Reaction Time in ms
None	1	2200
Self-Defined	0.67	2141
4 kHz Burst	1.24	2170

Table 7.2: Average deviation of the dialed-in value from the target value dependeing on the used adutory feedback

which shows, that pleasant auditory feedback increases both reaction time (and hence security) and accuracy (and hence user-friendliness) of a rotary encoder as an input device.

Chapter 8

Proposed New Method for the Prediction of the Acceptance of Acoustic Feedback

Since the perceived Quality of a sound is an impression generated by the reflective part of the human perception it is an impression which varies greatly from listener to listener. Furthermore the quality of a sound is context dependant i.e. a loud engine sound might be desirable in a sports car while it is disturbing in a luxury sedan. A distinct metallic click sound might be suitable and therefore perceived as high quality on a combination lock but might be regarded as a sign for damage in a computer hard drive and so on. These examples show that it is not feasible to design a method which can automatically identify which is good per se. For this reason the proposed method uses a subjectively selected reference sound. It is assumed that this reference sound is perceived as good by the subject, or at least as good as possible within practical limitations. This sound shall subsequently be used as a benchmark to compare arbitrary sounds against. It is important to understand that the sounds which are compared have to be of the same nature. This is self-explanatory for tests of other electromechanical feedback sounds but has to be taken into account is synthetic feedback sounds for i.e. touch screens are examined.

8.1 Outline

At the beginning of automatic feedback prediction the arbitrary reference sounds have to be analyzed first. The analysis results are stored as reference values for the subsequent comparison with sounds that have to be examined. The

8.2 Analysis Program "proBand"

The analysis program is called proBand which is a reference to both its intended use as a replacement of the extensive use of test persons in quality control and its working principle since the signal is split up in several frequency bands and then analyzed per band.

8.2.1 proband.m

The analysis program is implemented as a MATLAB function which uses the common wav-format as input, it is also possible to use the industry-standard hdf/dat-format by Head Acoustics. The program assumes a sample rate f_s of 96 kHz which is common with high quality recording equipment. If the sample rate of the input file differs, the input signal is resampled. The signal is then analyzed as a whole and subsequently for each frequency band. The proposed bands are the 3rd octave bands specified by DIN 61260 [?, ?, din] specifically the bands with center frequencies of 800 Hz and higher since the common control elements do not have relevant spectral components below this band. For this reason and since the measurement equipment does not provide very good ambient noise rejection below this reason, the signal is filtered using a FIR high pass filter with a cutoff frequency of 500 Hz before analysis. The resulting signal parameters are the same for both the analysis of the sum signal and the analysis of each band:

- Peak Amplitude
- RMS Amplitude
- Amplitude Envelope (leading to attack and decay time)

The calculations of the amplitude envelope and subsequently the attack and decay times take place in subroutines. The calculation however only takes place if the peak amplitude within a frequency band is at least 10 dB higher than the RMS amplitude. This criterion is required since actual recordings are subject to microphone and amplifier noise and the signal amplitudes are in general relatively small.

enveloper.m

This subroutine is called from the proband-program and uses a given input signal which can be either the high-pass- filtered sum signal or a band-pass-filtered signal. At the beginning of the routine the absolute peak value of the amplitude as well as its sample number are detected. Beginning from this absolute peak the signal the routine moves forward and backward through the dataset to and marks the local absolute maximum for the interval from the first sample to the current sample for the attack portion and from the current sample to the last sample for the decay portion. The result is a raw amplitude envelope as seen in *figure 8.1*.

The resulting amplitude envelope is subsequently returned to the main program where it will be used for the extraction of attack and decay times in other subroutines.

attacktime.m

First, the attack time is calculated. As mentioned before, the raw amplitude envelope is used as input for this subroutine. The peak value of the amplitude envelope and sample index of the peak value are used as reference. Furthermore, the mean value of the envelope is calculated. The routine starts at the sample index of the peak value and is then decremented until the value at the current sample index is either below the mean value or below 25 % of the peak value. The difference between this sample index and the index of the maximum

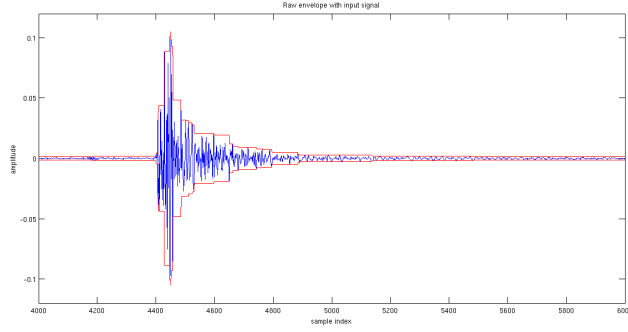


Figure 8.1: The blue line is the input signal, the red line the raw amplitude envelope as it is returned from `enveloper.m`, the image has been zoomed in to the actual click sound

is considered the attack time in samples which is returned to the main program after a conversion into seconds.

decaytime60db adaptiv.m

The subsequent calculation of the decay time is done in a similar way in this accordingly named subroutine. It is important to understand that the decay phase of the amplitude envelope is considered to be exponential. Like the attack time detection the routine first detects the peak value of the amplitude envelope as well as the sample index i of the maximum. Every sample before the detected maximum will be discarded for further processing, so the maximum will be located at sample index $i = 1$. Starting from this position the index is incremented until the value at the current index position is less than the mean value of the decay phase. This index shall be called i_{noise} . This point on which the measured signal disappears in the noise floor is sufficient to extract the logarithmic attack time. But as can be seen in figure [?] due to the exponential nature of the decay it can lead to significant errors in the identification of the decay time depending on the level of the noise floor.

To overcome this, the following approach shall be used:

The general form of an exponential decay is given as:

$$x = x_0 e^{-\delta t} \cos(\omega t + \varphi) \quad (8.1)$$

with $\cos(\omega t + \varphi)$ being the oscillation and $x_0 e^{-\delta t}$ being the amplitude envelope of this decay. x_0 is the initial amplitude. From what is known from the field of psychoacoustics, that the human sense of hearing acts as an integrator and is less sensitive to relatively quiet sounds following a relatively loud stimulus. It is therefor necessary that the integral of the modelled decay

$$\int_1^{n_{noise}} x_0 e^{-\delta t} dt \quad (8.2)$$

matches the integral of the raw decay which is taken from the measured sample. As the sample is available in time-discrete form only the, with $rawdecay(i)$ being the variable in which it is stored, the resulting equation is formulated as

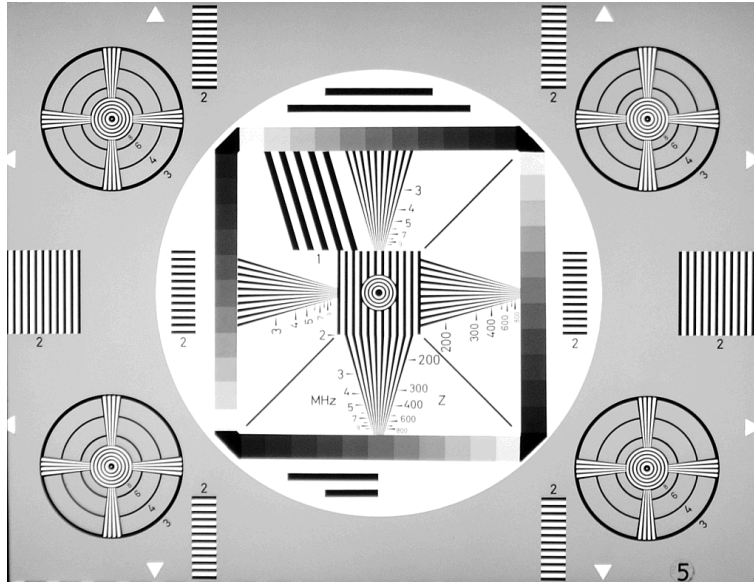


Figure 8.2: Noise Floor...man sollte hier sehen, wie verschieden hohe noise floors das Ergebnis beeinflussen koennen

$$\sum_{i=1}^{i_{noise}} rawdecay(i) = \sum_{i=1}^{i_{noise}} x_0 \cdot e^{i \cdot d} \quad (8.3)$$

Finally, the program returns the t_{60} time to the main program.

8.2.2 Tools

Some conversion tools have been written as a part of the software package which was used in this work.

damp 2 t60.m, t60 2 damp.m

This two programs convert damping coefficients into the respective 60dB decay time and vice versa.

db2pa.m, pa2db.m

This two programs convert the sound pressure from the logarithmic dB scale to the linear pascal scale and vice versa.

8.3 Correlation

After the analysis of the signals, the values are stored in a struct-array containing the following information:

- Work Directory
- Filename
- Attacktime per Band
- Decaytime per Band
- Peak Value per Band
- RMS Value per Band
- Damping Coefficient per Band
- Overall Signal Parameters
- Version of the Analysis Program

As justification by the overall amplitude is insufficient as explained above, the correlation coefficient of one or more signal parameters to a given reference are proposed as signal feature to judge the subjective acoustic quality of a stimulus. The idea is that stimuli which are similar to a *good* reference stimulus are likely to be perceived good as well while differing from the reference decreases the chance of being well accepted.

For the extraction of theoretically good sounds from the database a reference is needed. This can be acquired by one of the following three methods:

- Selection of an existing sound from the database
- Recording and subsequent analysis of a physical sample of a reference control element
- Design and subsequent analysis of a target sound using the synthesizer which is explained above

The computation of the correlation coefficients of one signal feature for 15 bands and all suitable recordings in the database can be performed within under 10 seconds on an average computer system using Matlab. Since the third subjective trial showed almost no impact of the attack time and since the damping coefficient and the decay time describe the same phenomenon, the correlation of the damping coefficients and signal peaks are sufficient to compare the signals to a reference.

8.4 Judgement / Selection

In the final step, signals are justified according to their match with the reference sound. This match is established by considering the correlation and - due to the high impact of the amplitude as shown in trials 1 and 3 - the overall amplitude of the signal.

Signals which exceed the desired peak amplitude by more than 10 dB are very likely to fail in a subjective comparison even if the spectral composition is very similar. On the other hand, signals which match the reference from an overall amplitude point of view might fail if the

spectrum is very *dull* if the preferred signal is supposed to be *bright* or vice versa. The actual impact of the bandwise correlation of the signal parameters and the suitability as a method of predicting the acceptance of the acoustic feedback of a rotary encoder shall be examined and proven in an experiment which is described in the next chapter.

8.5 Experimental Validation of the Predictor

This chapter describes the validation of the method for the prediction of the acoustic feedback of a rotary control element as described above. The basic idea is that each subject will design an individual reference sound which is fed into the analysis system. The analysis algorithm shall then select sounds from the database which are likely to be perceived well by the subject as well as sounds which are supposed to be rejected.

8.5.1 Design of Experiment

The experiment consists of two parts, separated by a short break in which the calculation is performed:

Part 1 - Target Sound Design

in the first part the subject's task is to design a feedback signal which is perceived pleasant by the subject. This is done using an approach similar to the approach in the fourth subjective trial [LNIK], leading to a relatively similar control panel (see *figure 8.3*).

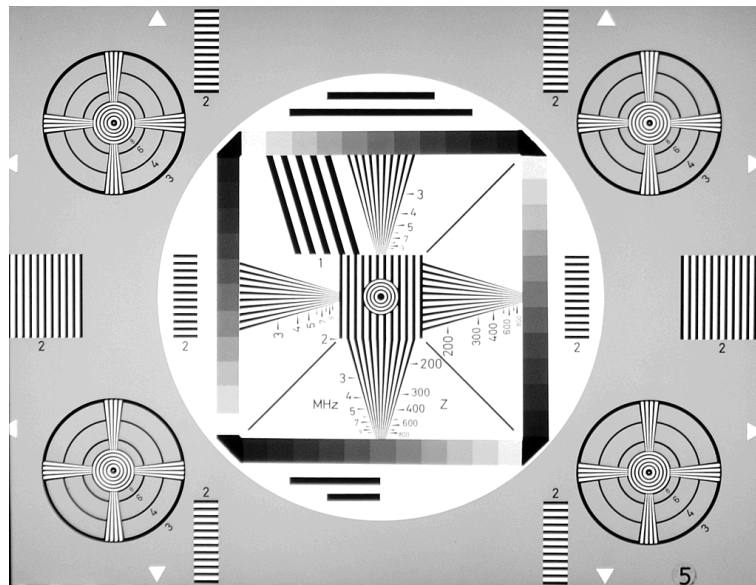


Figure 8.3: User interface for Sound design in the final experiment, the subject can rate the stimuli on a seven point scale ranging from extremely bad to extremely good

The only obvious modification from the interface which was used to design the signal before is that the spectrum parameter can be controlled in an adaptive way now, too. Before it was only possible to select the desired spectrum from a range of pre-defined spectra but because of the previous results a synthesis model has been conveyed which allows the modification of the spectrum on an abstract level.

However, the software behind the user interface was modified compared to the fourth series of trials. The implementation of the synthesizer in Matlab provided advantages over it's

implementation in Puredata regarding the handling of the acquired data. For this reason, Puredata only provided the user interface and the playback environment which has proven very stable and reliable regarding the timing.

As before, the subjects could continue to modify the sound even if the indicators for the target step-size showed that the sound has been set-up as accurately as desired. This is because of the fact that the parameters influence each other. Once the subject is satisfied with the sound, the confirm button stores the current settings and finishes this part of the experiment.

Calculation

In the calculation step the stimulus which was designed by the subject is analyzed and compared to the stimuli in the database using the methods explained in the previous chapter. The program selects a total of 29 sounds from the database, which are selected due to the following criteria:

- Increasing Correlation Coefficient of the Damping Coefficients (10 stimuli)
- Increasing Correlation Coefficient of the Peak Amplitudes (10 stimuli)
- Match of the Overall Amplitude
 - 10 dB too quiet (3 stimuli)
 - match (3 stimuli)
 - 10 dB too loud (3 stimuli)

Since the sounds which are stored in the database are all relatively similar due to the fact that they are all based on physical control elements which are constructed in a similar way, their typical correlation coefficients were similar as well. *Figure [?]* shows an example of how the typical distribution of correlation coefficients looks like:

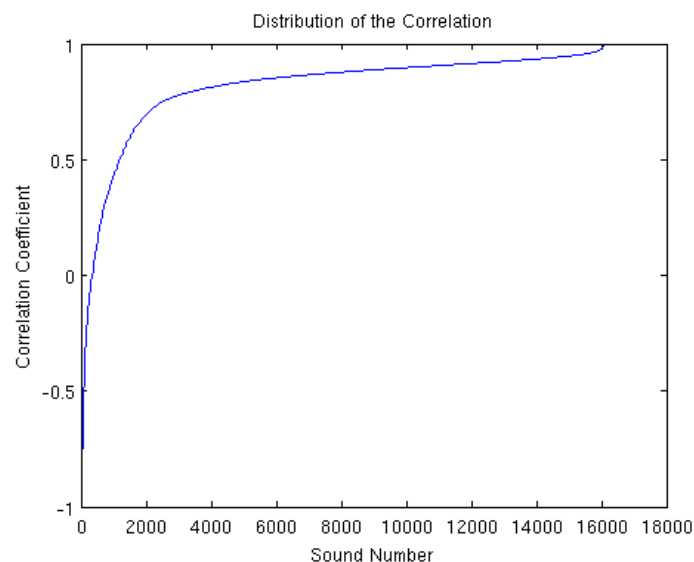


Figure 8.4: Note that over 75% of the sounds feature very high correlation coefficients

The figure shows the correlation coefficients for the peak amplitude parameter for the entire database. It can be seen that the curve is relatively steep for the first 2000 sounds and then becomes relatively flat, more than 75% of the sounds show high to very high correlation with the reference value in this example. The same is true if the damping coefficients are correlated.

The analysis and subsequent correlation and data extraction can be done in less than a minute so that the subjects can continue with the second part of the experiment right away, since the time used for calculation is used to brief the probands with their task for the second part.

Part 2 - Evaluation

In this part of the experiment it is the subject's task to rate the sounds which have been selected by the system in the step before. The subjects do this using a seven-point scale on the interface which is shown in the following *figure 8.5*:

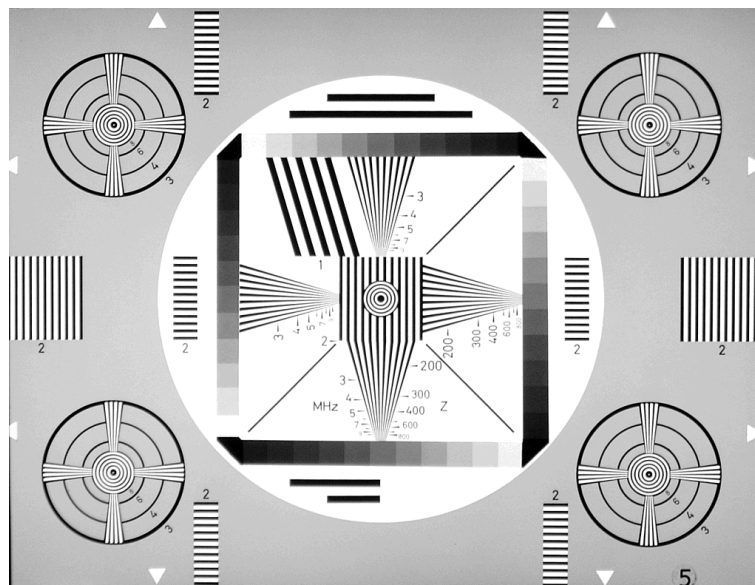


Figure 8.5: Example for task in the third test stage

The subjects could listen to the sound as long as they wished. Once the subject came up with a conclusion, he could enter the rating of the sound into the seven-point scale and confirm the result. The next stimulus would be played back using the simulator.

The order in which the stimuli were presented to the subjects was randomized for every subject. This was done by writing the task number, the sound index and a random number into a 4x29 matrix, which is partially shown in the following *table 8.1* .

Task Index	Sound Index	Random	Rating
1	2982	0.552	
2	12	0.022	
3	5298	0.921	
4	15229	0.731	
....	
29	1223	0.001	

Table 8.1: Example for task in the third test stage

This table is then sorted according to the "Random" column. The subjects' answers are stored in the "Rating" column. After the end of the test, the matrix can be sorted using the "Task Index" column again, discarding the "Random" column.

The 20 sounds which are selected because of their correlation coefficients are normalized in terms of peak amplitude in order to prevent any influence through this parameter.

8.5.2 Analysis of the Results

The results of both the individually optimized feedback signals and the validation of the prediction algorithm will be examined in this section. The First provides further insight on the kind of feedback signals the subjects desire. The latter shall prove that the proposed methods for the improvement of the prediction of acceptance of acoustic feedback are valid.

Part 1 - Individually optimized feedback signals

After the subjects have configured and confirmed their individually desired feedback signal, these signals are subsequently analyzed and compared. The attack time of the signals was fixed, what they could adjust was the slope of the spectrum (i.e. emphasis of high frequencies, uniform distribution or emphasis of low frequencies), the overall peak amplitude and the damping coefficients.

The subjects typically preferred uniform spectra or spectra with over-emphasized high frequencies. *Figure 8.6* shows the individual results of the preferred spectra.

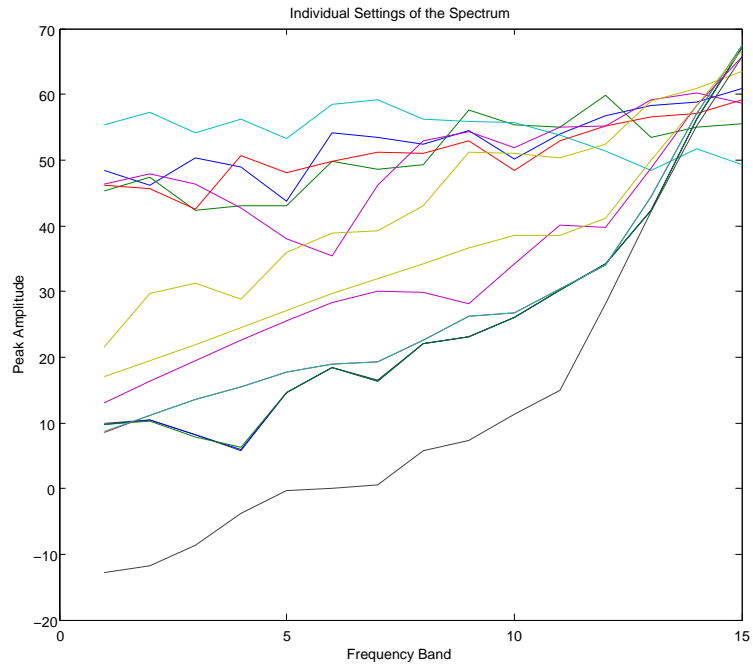


Figure 8.6: The individually set-up spectra show that the subjects prefer a certain over-emphasis of high frequencies

The figure shows that some probands emphasized high frequencies very much while no subject emphasized the low frequencies. This is interesting because of the free-text answers which the subjects from the first trial gave. Adjectives like dull were relatively common among attributes for good sound while bright was very often considered to be bad.

The following *figure 8.7* shows the setting of the individually optimized damping coefficients.

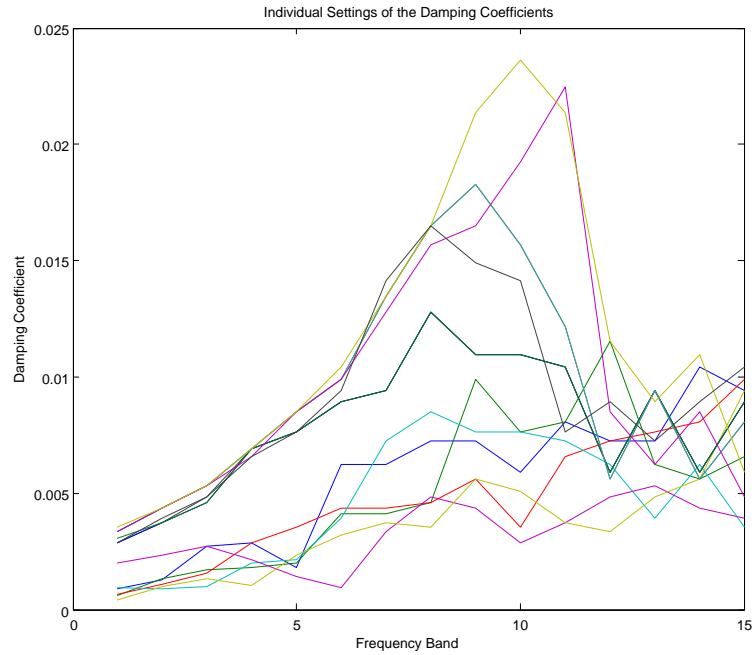


Figure 8.7: Individual settings for the damping coefficients per band. The majority of subjects adjusted the damping parameter in a similar way

The damping coefficients for lower frequencies are lower than for the higher bands for all subjects. This is already implemented in the synthesis algorithm since mechanical oscillations typically show higher damping coefficients and hence shorter decay times at higher frequencies. The subjects could only affect the damping coefficient as a whole. However, the change for higher frequencies was larger than at lower frequencies. Nevertheless, the majority of subjects adjusted the damping parameter to the same range.

Finally, *figure 8.8* shows the individual settings for the peak amplitude:

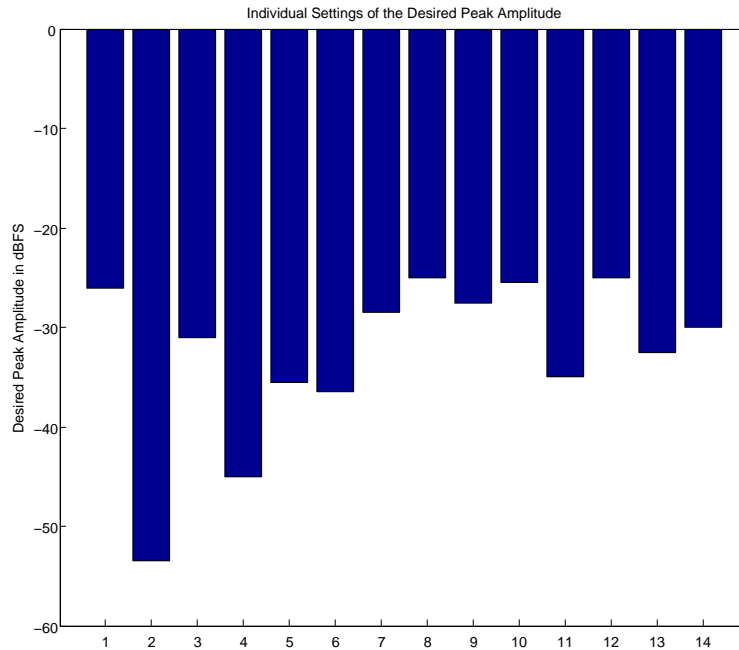


Figure 8.8: Individual desired peak amplitudes in dBFS, note that the majority of answers is within a 10 dB range.

This result supports the results from the previous experiment 4, since the majority of the subjects adjusted the desired amplitude to values within a 10 dB range, while only 2 of 14 subjects desired significantly lower values.

Part 2 - Rating of automatically selected sounds

The acceptance has been predicted based on three different methods:

- Correlation of the Damping Coefficients
- Correlation of the Peak Amplitudes
- Match of the Overall Peak Amplitude

The proposals for a prediction method which are based on correlation of signal parameters are supposed to have an effect on the acceptance of the stimuli, i.e. high correlation coefficients tend to higher acceptance ratings than low correlation values. A comparison of both methods of selection is shown in *figure 8.9* :

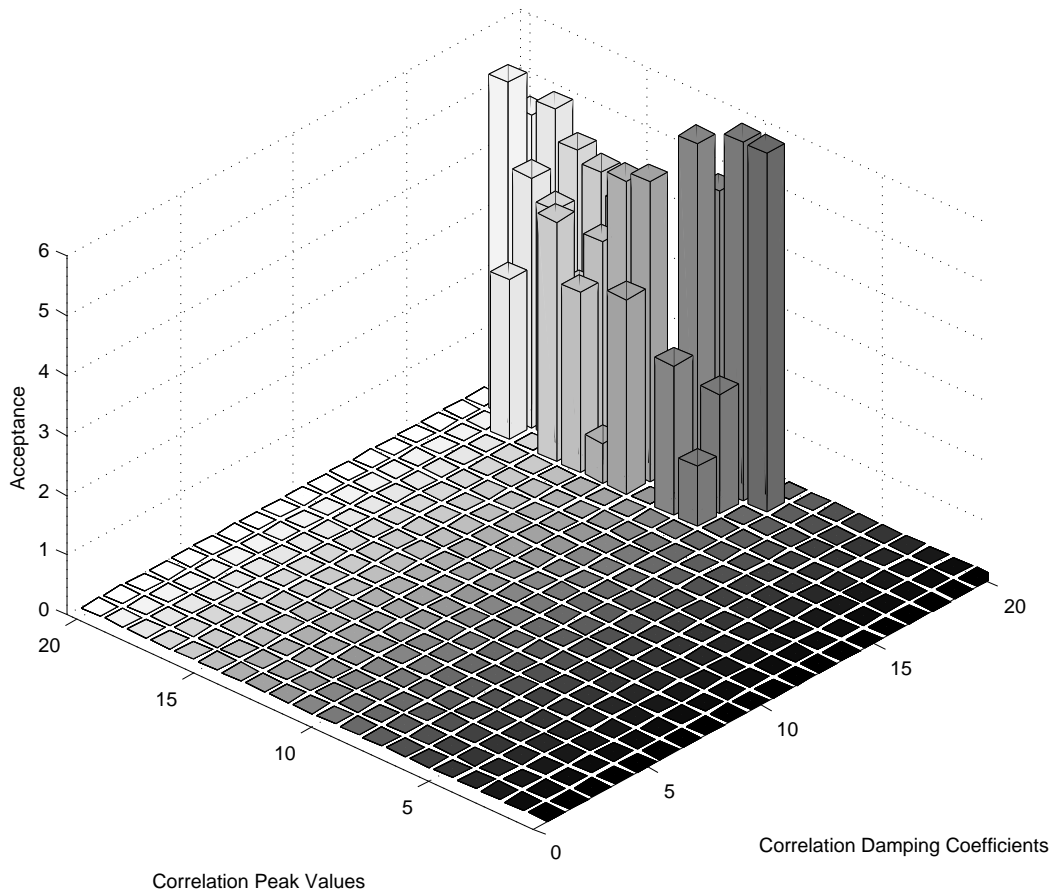


Figure 8.9: This plot shows the average acceptance values achieved with stimuli which showed the accordings combinations of the correlation coefficients of peak amplitudes per band and damping coefficients per band.

Since the variations in the actual correlation coefficients which have been calculated for each sound are very small the values have been quantized into 20 ranges and averaged within each range. The result is shown in the above figure. x- and y-axes denote the correlation coefficients of the damping coefficients and the peak amplitudes respectively. The z-axis denotes the average acceptance within a certain correlation coefficient bin. One would expect that the highest z-values occur in the upper ends of both x- and y-axis.

This is true for the correlation coefficient of the damping coefficients, in which relatively small deviations from the highest possible values lead to a significant reduction in the acceptance values. Regarding the correlation coefficients of the peak values it can be seen that even relatively low values lead to high acceptance ratings as long as the damping coefficients correlate well with the reference.

A possible explanation for this is that the human sense of hearing works as an integrator and as long as the damping coefficients correlate well, the integral of emitted sound per band should correlate relatively well, too, even if the peak amplitude in this band did not match the desired values.

As far as the influence of the absolute peak value on the acceptance is concerned, the results from the very first subjective experiment remain valid, as can be seen in *figure 8.10*

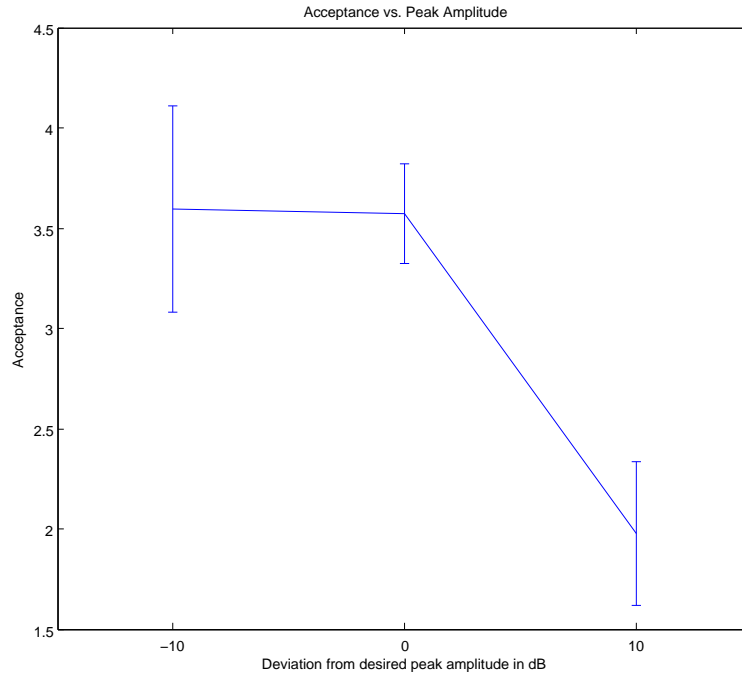


Figure 8.10: This figure shows that exceeding the desired amplitude leads to a significant decrease of the average acceptance of the sounds

Just like in the first subjective trial, the stimuli which exceed the desired amplitude receive significantly worse acceptance values. While the acceptance with the individually desired peak amplitude lead to a very similar acceptance value than the stimuli which were 10 dB too quiet, the standard deviation of the results is smaller with the stimuli which match the desired peak amplitude.

8.5.3 Conclusion

The experiment has shown that at least one of the two proposed methods for the improvement of the prediction of the acoustic feedback in fact leads to higher acceptance values in the subjectively rated sounds.

Chapter 9

Results

This section sums up what has been achieved withing this project and formulates the improvements to the topics which habe been pointed out in *chapter 4*.

9.1 Method for Acoustic Rapid Prototyping

At first, a new methodology (see *figure 9.1*) for the definition of target sounds has been developed and successfully tested.

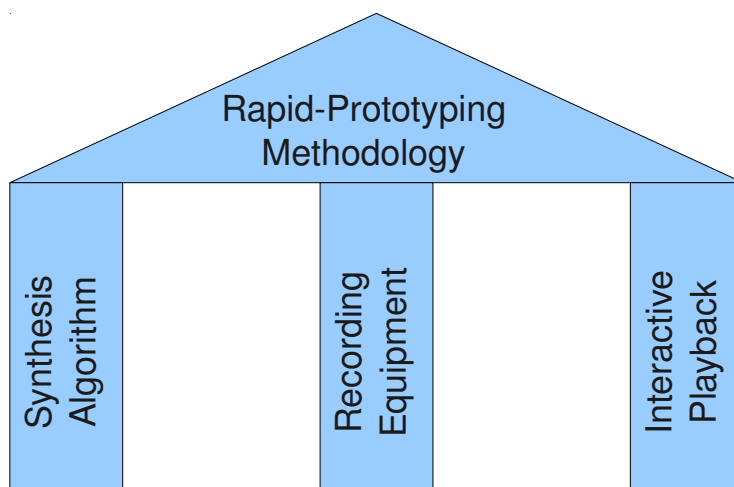


Figure 9.1: Visualization of the acoustic rapid prototyping mehtodology which uses synthesis and interactive playback as its foundataions.

Basically the method consists of an adaptive approach to the target sound. This has been made possible by reduction of the parameters which are needed to synthesize a realistically sounding acoustic feedback signal. The final significant enhancement to traditional sound design approaches is the acoustic simulator. This device enables the designer of the sound to experience the sound in it's natural context. Optionally, the method can take measurements

of physically existing control elements into account and use them as the basis for future sound design steps.

9.2 New Algorithm for Prediction of Acoustic Acceptance

Based on the results which were gathered in extensive subjective trials, an new algorithm for the prediction of the acceptance of the acoustic feedback of a control element has been formulated. The new idea of the proposed method is to acknowledge that the acoustic feedback cannot be predicted on an absolute scale but it context sensitive and subjective. This leads to the conclusion, that the reference which is used as a standard for the prediction has to be context sensitive and subjective as well.

However, this insight in combination with the method for rapid prototyping which has been introduced in this work provides a possibility to objectively predict the acceptance of a sound within a given context. The extraction of the relevant acoustic features - which are the same as for synthesizing a convincing acoustic feedback stimulus - and the subsequent correlation with a reference has proven successful. The reference can be obtained from applying the introduced prototyping method to either a small number of experts or a larger number of naïve listeners and subsequent averaging of the obtained individual results.

9.3 New Proof for the Effects of Acoustic Feedback

Finally, the conducted subjective trials showed that the acoustic feedback influences the performance of subjects. This is both true from a speed and from an accuracy point of view. While the increases in speed are relatively small compared to the overall duration of an operation they are certainly not to be neglected in the context of a car or in the operation of critical machinery. The improvement in accuracy however is significant and certainly shows that acoustic feedback can in fact improve usability, especially in a context in which there are now haptic cues.

Chapter 10

Conclusion

Finally it has to be said, that during the project a number of interesting questions has arisen which could not be dealt with in the given time but provide challenges for further research for which this work can hopefully serve as a good starting point.

As the quoted works of Brewster [LINK], Altinsoy [LINK] and Reisinger [LINK] show, the feeling of the feedback of control elements, whether it is provided acoustically, visually or tactile, gained importance in academic research and is still gaining importance in a number of applications.

Especially the indermodal effects which occur when several ways of feedback are used simultaneously are as of now very little understood and provide further opportunities.

But even the isolated examination of the effects of acoustic feedback still leaves room for further examination. This work focussed exclusively on the feedback of physical control elements and hence the developed models are exclusively useful to evaluate or synthesize signals like they occur in physical control elements. The success of touch-screen operated mobile phones or navigation devices in recent years as well as the advent of tablet PCs and new developments like camera-observed gesture control show that many appliances in the future are likely to be operated not my pyhsical control elements anymore but by virtual input devices in some way or another. Adding acoustic feedback to such a virtual input device always requires the playback of synthetic sounds. Depending on the context, results of this work may be used as design tools.

Furthermore, the playback of synthetic acoustic feedback enables the designer of the appliance to adjust the feedback signal depending on the current conditions. In a car environment this means that a signal might be very quiet when the car is parked, but significantly louder and high speed and hence noise.

Furthermore, the usefulness of acoustic feedback has currently only be proven in a laboratory environment with an abstract task. Future work might deal with the question, how much the awareness of the driver is increased in for example an driving simulator. However, tools and methods developed in this work can be transferred for this purpose.

In general, one can state that the increasing complexity of today's technical products, the icreased demand for comfort as well as the demographic change will set the future challenges in user interface design, and the concise and well planned use of acoustic cues can provide improvements for both usability and user experience in this field.

Appendix A

Data Sheets

This section contains data sheets of the most important components of the measurement system and the acoustic simulator as well as the encoder which this work focusses on.

In detail the datasheets are:

Greyhill 62AG - Optomechanical Encoder for human machine interfaces

Hypex UCD-180HG - Class-D amplifier module used in the acoustic simulator

Hypex SMPS-180 - Switching mode power supply unit for the amplifier module

Greyhill Series K63 - High-resolution encoder used in the acoustic simulator

Scanspeak Revelator 1" Tweeter - Loudspeaker of the improved acoustic simulator

Mackie M-1400 - Amplifier for the stepper motor of the measurement system

DIN EN 61260 - Standard for Third-octave and octave band filters

SERIES 62AG

Price Competitive Solution



FEATURES

- Long Lasting (1 million cycles)
- Optional pushbutton
- Available in 16 and 32 Detent Positions
- 4 inch cable / connector assembly

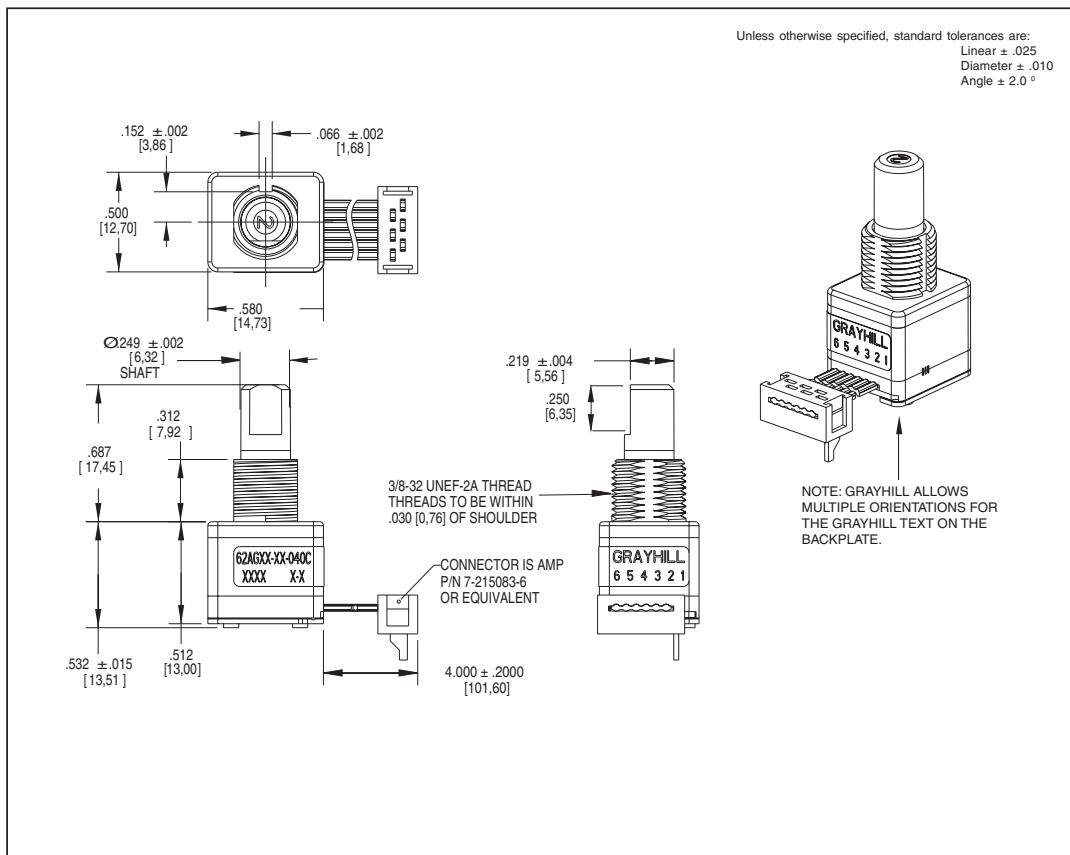
APPLICATIONS

- Automotive audio, navigation & driver information systems
- Medical Equipment
- Test & Measurement Equipment
- Audio & Video Equipment

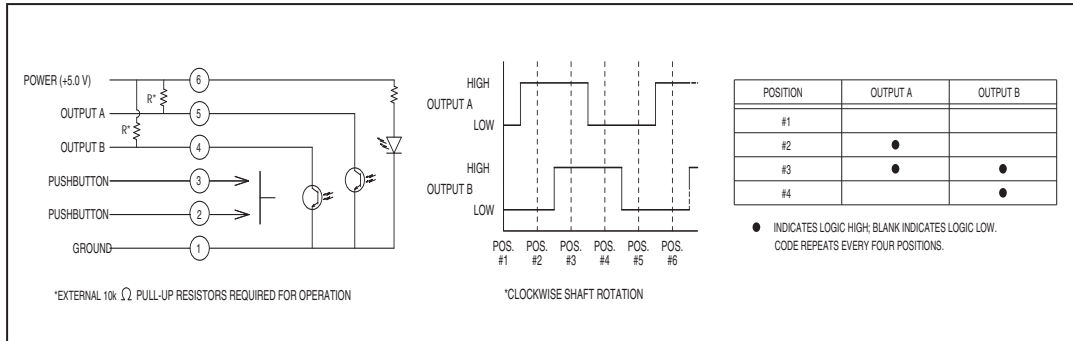


DIMENSIONS In inches (and millimeters)

Optical and Mechanical Encoders



WAVEFORM AND TRUTH TABLE



SPECIFICATIONS

Environmental Specifications

Operating Temperature Range: -40°C to 85°C

Storage Temperature: -43°C to 38°C

Humidity: 96 Hours at 90-95% humidity at 40°C

Mechanical Vibration: Harmonic motion with amplitude of 15g within a varied frequency of 10 to 2000 Hz for 12 hours Mechanical Shock

Test 1: 100g for 6 ms half-sine wave with a velocity change of 12.3 ft/s.

Test 2: 100g for 6 ms sawtooth wave with a velocity change of 9.7 ft/s.

Rotary Electrical and Mechanical Specifications

Operating Voltage: 5.00±0.25 Vdc

Supply Current: 30 mA maximum at 5 Vdc.

Logic Output Characteristics:

Logic high shall be no less than 3.0 Vdc Logic low shall be no greater than 1.0 Vdc

Minimum sink current: 0.5 mA for 5 Vdc.

(Preliminary)

Power Consumption: 150 mW maximum for 5 Vdc

Output: Open Collector Phototransistor

Optical Rise Time: 30ms maximum.

Optical Fall Time: 30ms maximum.

Average Rotational Torque:

2.0±1.4 in-oz before life. 50% of initial value after 1 million cycles.

Mechanical Life: 1,000,000 cycles of operation. 1 cycle is a rotation through all positions and a full return.

Mounting Torque: 15in.-lbs. maximum

Shaft Pushout Force: 45 lbs. minimum

Terminal Strength: 15 lbs. Cable pull out force minimum

Solderability: 95% free of pin holes and voids

Maximum rotational speed: 100 rpm.

Pushbutton Electrical and Mechanical Specifications

Rating: 10 mA @ 5 Vdc

Contact Resistance: <10 W (Compatible with CMOS or TTL)

Life: 1 million actuations minimum

Contact Bounce: <4 ms make, <10ms break

Actuation Force: 510±150 grams

Shaft Travel: .017 ± .008 INCH

Materials and Finishes

Bushing: Zamak 2

Shaft: Zamak 2

Detent Rotor: Reinforced Nylon Zytel 70G33L UL 94

Detent Spring: 303 Stainless Steel

Housing, Upper: Nylon 6/6 25% glass reinforced. Zytel FR-50

Light Pipe: Lexan, GE

Code Rotor: Delrin 100

Housing, Lower: Nylon 6/6 25% glass reinforced. Zytel FR-50

Pushbutton Actuator: Reinforced nylon. Zytel 70G33L UL 94

Pushbutton Dome: Stainless Steel

Printed Circuit Board: NEMA Grade FR4, Double clad with copper, Plated with gold over nickel

Infrared Emitting Diode: Gallium Arsenide

Phototransistor Diode: NPN Silicon

Resistor: Metal oxide on ceramic substrate

Spacer: Pet plastic

Backplate: Stainless Steel

Label: TT406 thermal transfer cast film.

Solder: 96.5% tin / 3% silver / 0.5% copper.

No clean.

Hex Nut: Brass, Plated with nickel

Lockwasher: Stainless steel

Cable: Copper Stranded with topcoat in PVC insulation

Connector (.050 center): PA4.6 with tin/nickel plated phosphor bronze.

Optical and Mechanical Encoders

Series
Angle of Throw: 22 = 22.5° for code change and 16 detent positions

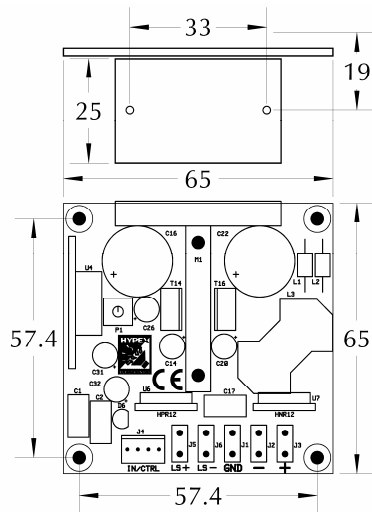
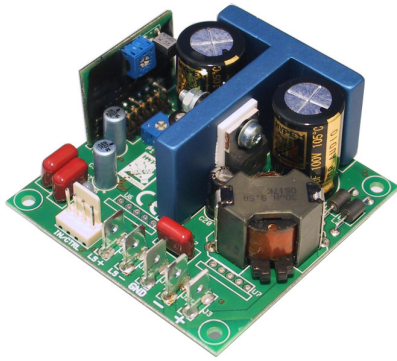
Pushbutton Option: 01=No Pushbutton, 02=With Pushbutton

Termination: C = .050 Center ribbon Cable with connector
Cable Termination: 040=4.0 inches. Cable is terminated with Amp Connector P/N215083-6.
See Amp Mateability Guide for Mating Connector for details.

62AG22-XX-040C

Available from your local Grayhill Component Distributor. For pricing an discounts, contact a local Sales Office, an authorized local Distributor, or Grayhill.

High Grade Audio Power Amplifier Module



Highlights

- Flat, fully load-independent frequency response
- Low output impedance
- Very low, frequency-independent THD
- Very low noise
- Fully passive loop control
- Consistent top performer in listening trials

Features

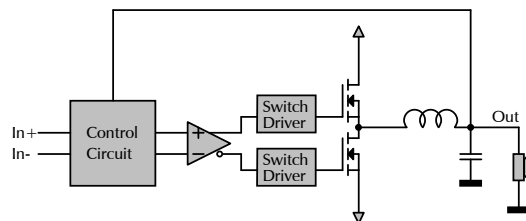
- Runs on unregulated +/- rails
- Pop-free start and stop control
- Differential audio input
- No compromise components
- LM4562 buffer OpAmp
- HxR12 ready
- Improved on-board buffer supply
- Overcurrent and overvoltage protection
- Weight: 90gms (3.1oz.)

Applications

- Monitor loudspeakers for recording and mastering studios
- Audiophile power amplifiers for professional and consumer use
- Public Address systems
- Home theatre systems
- Active loudspeakers

Description

The UcD180HG™ amplifier module is a self-contained high-performance class D amplifier intended for a wide range of audio applications, ranging from Public Address systems to ultrahigh-fidelity replay systems for studio and home use. Chief distinguishing features are flat frequency response irrespective of load impedance, nearly frequency-independent distortion behaviour and very low radiated and conducted EMI. Control is based on a phase-shift controlled self-oscillating loop taking feedback only at the speaker output.



Performance data

Power supply = +/-45V, Load=4Ω, MBW=40kHz, unless otherwise noted

Item	Symbol	Min	Typ	Max	Unit	Notes
Output Power	P_R	180	-	-	W	THD=1%
Distortion	THD+N	-	0.1	0.15	%	20Hz<f<20kHz. Pout<P _R /2
		-	0.008	0.01	%	20Hz<f<20kHz. Pout=1W
Output noise	U_N	-	30μ	35μ	V	Unwtd, 20Hz-20kHz
Output Impedance	Z_{OUT}	-	-	20m	Ω	f<1kHz
		-	-	150m	Ω	f<20kHz
Power Bandwidth	PBW		20-35k		Hz	
Frequency Response		10	-	50k	Hz	+0/-3dB. All loads
Voltage Gain	A_v	25.5	26	26.5	dB	
Supply Ripple Rejection	PSRR		65		dB	Either rail, all frequencies
Efficiency	η		92		%	Full power
Idle Losses	P_n		4		W	
Standby Current	I_{STRY}		10m		A	
Current Limit			10		A	Stop mode after limiting 40ms.

Absolute maximum ratings

Correct operation at these limits is not guaranteed. Operation beyond these limits may result in irreversible damage

Item	Symbol	Rating	Unit	Notes
Power supply voltage	V_R	+/-50	V	Unit shuts down when either rail exceeds 52V
Peak output current	$I_{OUT.P}$	10	A	Unit current-limits at 10A
Input voltage	V_{IN}	+/-12	V	Either input referred to ground
Air Temperature	T_{AMB}	65	°C	
Heat-sink temperature	T_{SINK}	90	°C	User to select heat sink to insure this condition under most adverse use case

Recommended Operating Conditions

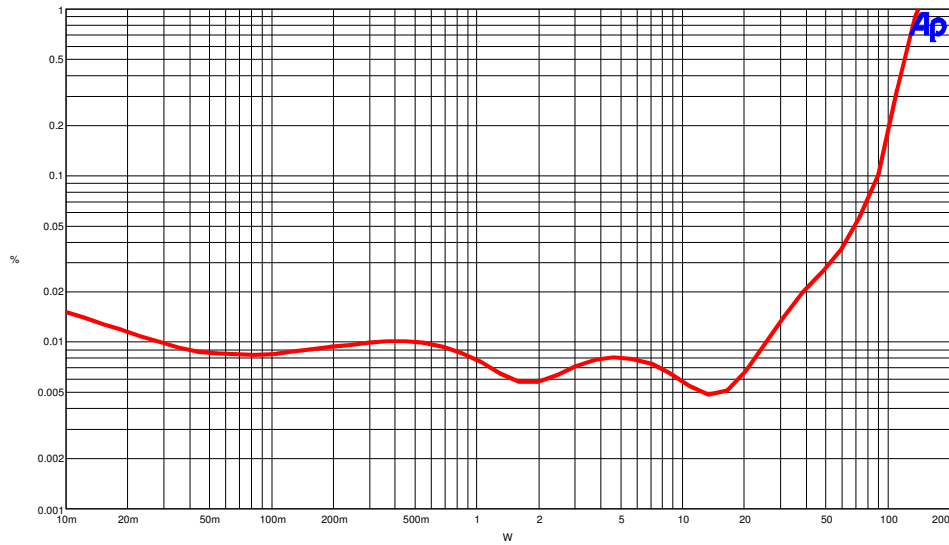
Item	Symbol	Min	Typ	Max	Unit	Notes
Power supply voltage	V_R	30 ¹⁾	45	50 ²⁾	V	
Load impedance	Z_{LOAD}	1			Ω	
Source impedance	Z_{SRC}			7k	Ω	Differential. Corresponds to 3dB noise increase.
Effective power supply storage capacitance	C_{SUP}	4700μ			F	Per rail, per attached amplifier. 4Ω load presumed.

¹⁾Unit shuts down when either rail drops below 25V.

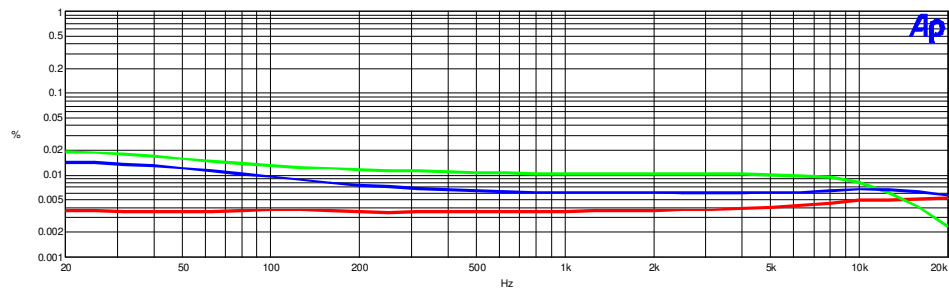
²⁾Unit shuts down when either rail exceeds 52V.

Typical Performance Graphs

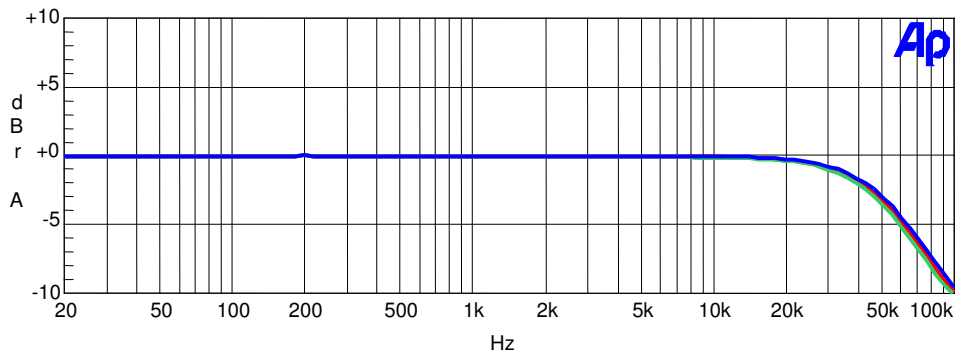
THD vs. Power (1kHz, 4Ω)



THD vs. Frequency (8Ω)

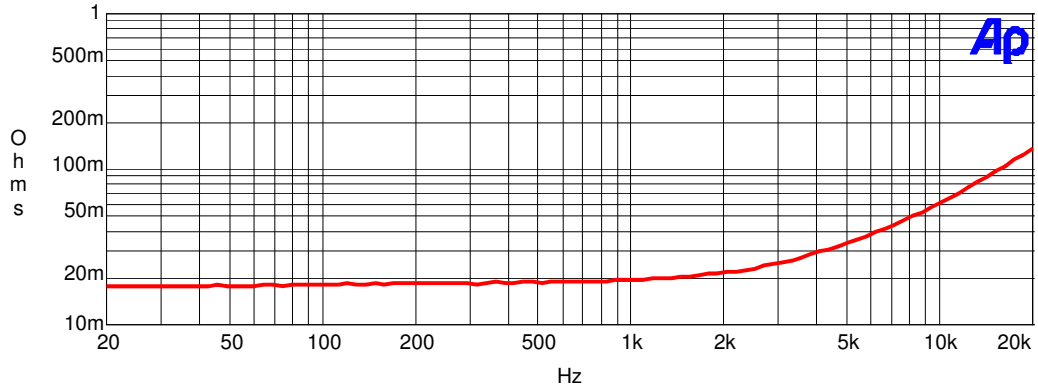


From top to bottom: 40W, 10W, 1W Frequency Response (4Ω, 8Ω and open circuit)

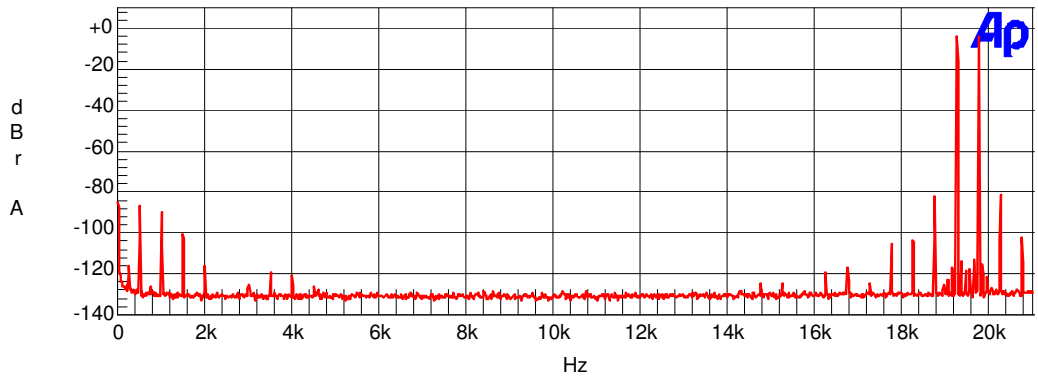


From top to bottom: open circuit, 8Ω, 4Ω

Output Impedance



19+20kHz IMD (10W, 4 ohms)



Universal Mains Input Audio SMPS



Highlights

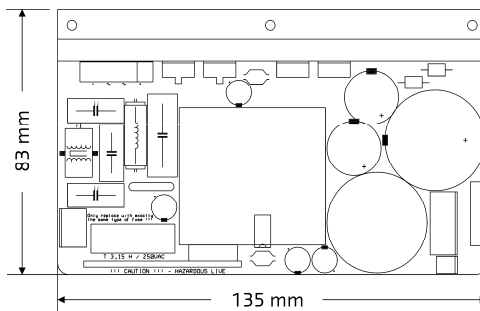
- High efficiency
- Universal mains input voltage
- Near unity Power Factor
- Low EMI

Features

- Advanced over current protection
- Remote controlled operation
- Low weight: 400 gr.
- Compact: 135 x 83 x 49mm
- Adjustable output voltage

Applications

- Supply for single or multiple amplifiers of the UcD™ range
- Active loudspeakers



Description

The SMPS180 is a high efficiency Safety Class 2 switch mode power supply specifically designed for use with our range of UcD™ amplifier modules. Key features are high efficiency over the entire load range, universal input voltage range (85 – 264V, 50 - 60Hz), near unity Power Factor, low weight and very low radiated and conducted EMI. The SMPS180 also features an advanced overcurrent protection which in case of temporary overload simply reduces the output voltage, only when the overload condition remains for a longer time the supply will enter hiccup mode until the overload condition disappears. This feature in combination with the advanced feedback topology and large secondary bulk storage output caps leads to the capability of delivering high dynamic headroom power to the connected amplifier. The SMPS180 also includes an auxiliary isolated $\pm 12V$ supply and a control circuit directly interfacing with our range of (OEM and standard) UcD™ amplifier modules. The supply is triggered for normal operation or latched off in case of critical fault via in built-in actuators. The SMPS180 is optimized from the first phase of design to final implementation to realize the low EMI signature required of the most demanding audio applications.

Absolute maximum ratings

Correct operation at these limits is not guaranteed. Operation beyond these limits may result in irreversible damage

Item	Symbol	Rating	Unit	Notes
Input voltage	V_{LINE}	264	Vac	
Air Temperature	T_{AMB}	50	°C	Power output is reduced
Heat-sink temperature	T_{SINK}	95	°C	

Recommended Operating Conditions

Item	Symbol	Min	Typ	Max	Unit	Notes
Operating Line Input Voltage	V_B	85		264	Vac	
Full Power Operating Input Voltage	$V_{B,FP}$	100		264	Vac	

General Performance data

Item	Symbol	Min	Typ	Max	Unit	Notes
Output Voltage	V_{OUT}	2 x 35		2 x 45	Vdc	See Note 4
Output Current	I_{OUT}		3.4		A	See Note 3
Regulated Output Voltage Aux	$V_{OUT,AUX,REG}$		2 x 12		Vdc	
Unregulated Output Voltage Aux	$V_{OUT,AUX,UNREG}$	2 x 14		2 x 18	Vdc	See Note 5
Output Current Aux	$I_{OUT,AUX}$		250m		A	per rail
Output Power	P_R	300	-	-	W	See Note 1
Audio Output Power @ 20Hz into amplifier load	P_{RALF}	180	-	-	W	See Note 2
Efficiency	η		TBD		%	full power
Idle Losses	P_0		2		W	
Standby Power	$P_{standby}$		TBD		W	
Switching frequency	F_{SW}		100		kHz	
Line regulation			TBD			%
Load regulation			TBD			%

Note 1: Output Power delivered to a resistive dummy load (generally the only specification supplied by other SMPS manufacturers).

Note 2: An audio amplifier actually draws twice the RMS power from the power supply. At high frequencies the secondary storage output caps are capable to provide this power. At very low frequencies however the SMPS is responsible for delivering this peak power to the amplifier.

Note 3: Both rails loaded. Maximum current per rail is 6.7A (one rail loaded).

Note 4: Adjustable by means of a potentiometer.

Note 5: This output voltage is proportional to main outputs Vcc and Vee respectively.

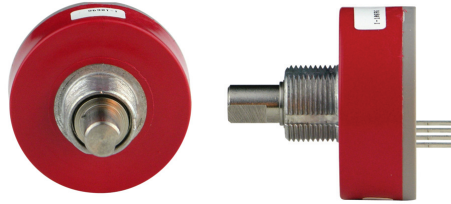
SERIES 63K

High Resolution, Ball Bearing,
4-Pin

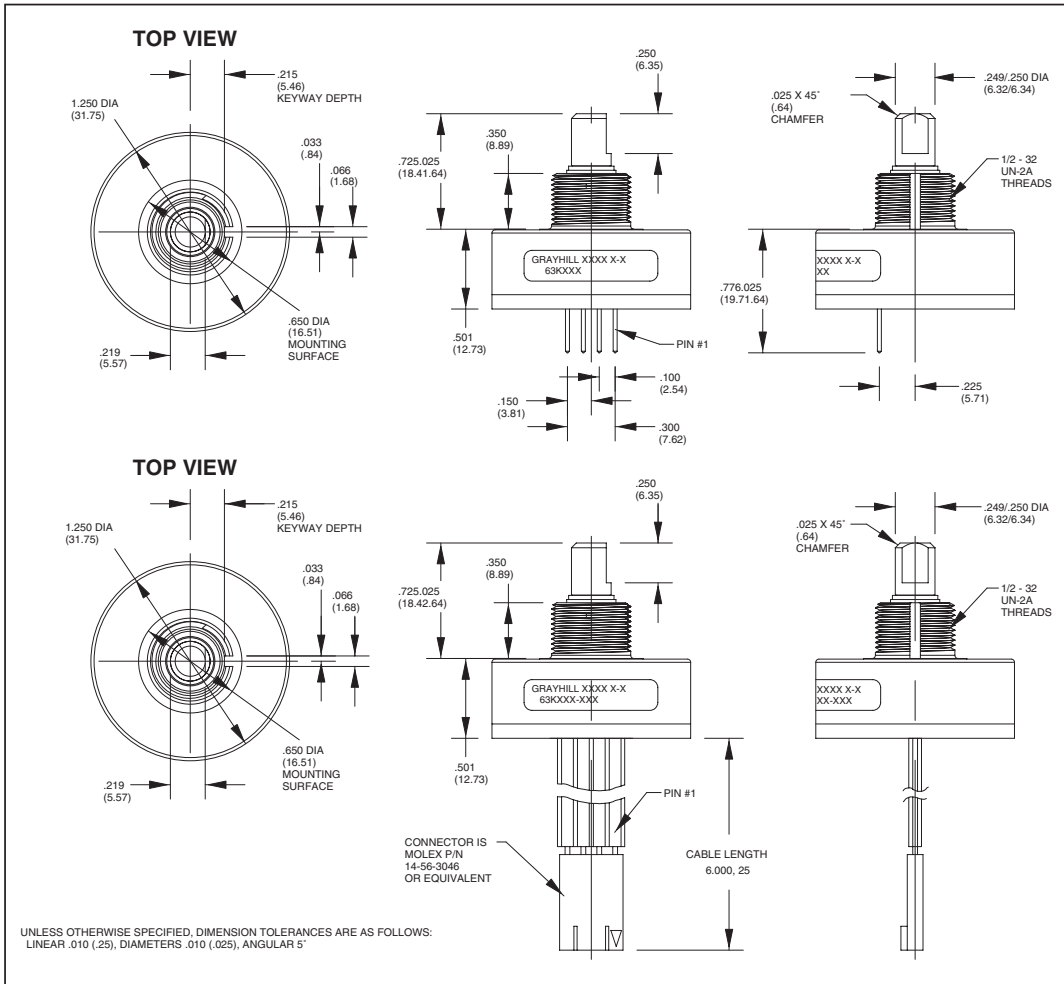


FEATURES

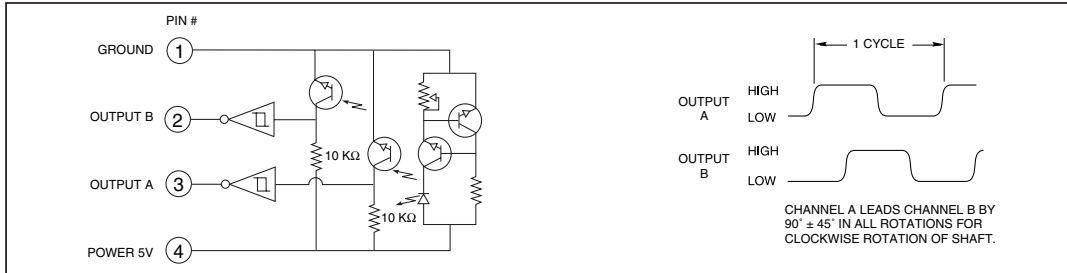
- 25, 32, 50, 64, 100, 128 and 256 Cycles per Revolution Available
- Sealed Version Available
- Rugged Construction
- Cable or Pin Version
- 300 Million Rotational Cycles
- 5,000 RPM Shaft Rotation



DIMENSIONS In inches (and millimeters)



Optical and Mechanical Encoders

CIRCUITRY AND WAVEFORM: Standard Quadrature 2-Bit Code

SPECIFICATIONS
Electrical Ratings

Operating Voltage: 5.0 \pm .25 Vdc
Supply Current: 30 mA maximum at 5 Vdc
Logic Output Characteristics:
 Output Type: Open collector with integrated Schmitt Trigger and 10 K Ω pull-up resistor
 Maximum Sink Current: 16 mA at .40 volts
Power Consumption: 150 mW maximum
Optical Rise Time: 500 nS typical
Optical Fall Time: 14 nS typical

Mechanical Ratings

Mechanical Life: 300 million revolutions
Time Life: Guaranteed for 10 years of continuous operation (calculated from emitter degradation data)
Mounting Torque: 20 in-lbs maximum
Terminal Strength: 5 lbs terminal pull-out force minimum
Solderability: 95% free of pin holes and voids
Operating Torque: 0.5 in-oz maximum (no detents) for unsealed versions
Externally Applied Shaft Force: Axial: 15 lbs maximum; Radial: 15 lbs maximum

Environmental Ratings

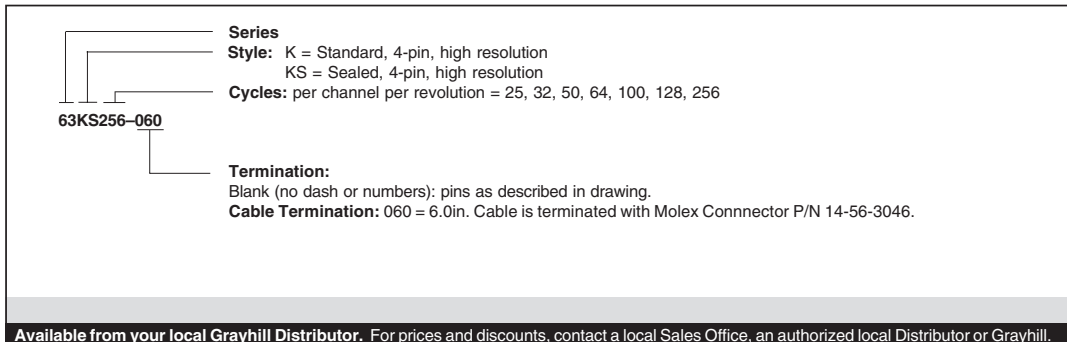
Operating Temperature Range: -40°C to 85°C
Storage Temperature Range: -55°C to 100°C
Relative Humidity: 90-95% at 40°C for 96 hours
Vibration Resistance: Harmonic motion with amplitude of 15g, within a varied 10 to 2000 Hz frequency for 12 hours per MIL-STD-202, Method 204
Shock Resistance: Test 1: 100g for 6 mS, half-sine wave with velocity change of 12.3 ft/s. Test 2: 100g for 6 mS, sawtooth wave with velocity change of 9.7 ft/s.

Materials and Finishes

Bushing: Zinc diecast
Housing: Hiloy 610B
Shaft: Stainless Steel
Code Rotor and Aperture: Chemically etched stainless steel/electroformed nickel
Printed Circuit Board: NEMA Grade FR-4. Five microinches minimum gold over 100 microinches minimum nickel over copper
Optical Barrier: Polyphenylene sulfide, 94 V-0
Backplate: Polyester
Header: Phosphor bronze, 200 microinches tin over 50 microinches nickel (pin version only)
Infrared Emitter: Gallium aluminum arsenide
Photo IC: Planar silicon
Retaining Ring: Stainless steel
Cable: 26 AWG, stranded/tinned wire, PVC coated on .100 (2.54) centers (cable version only)
Connector: Glass-filled PCT, UL94V-0

Bearing Subassembly

Bearing: NSK ABEC 5 (stainless steel)
Preload Collar: 303 (stainless steel)

ORDERING INFORMATION




**Revelator
1" Tweeter**

Type Number: D2904/980000

Features:

The Revelator series has for years been celebrated for producing the best sounding electro dynamic transducers in the world. Since ScanSpeak was founded in 1970, the audio engineers and R&D experts working on the line have been on a quest to create drivers that reveal all the sound in recordings, hiding nothing from the listener. This quest has resulted in several revolutionary inventions that remove distortion in the magnet systems and in the moving parts of the speaker. The philosophy is that the sound has to be very dynamic, giving a perfect transient response and providing tonal balance.



In the Revelator tweeters, additional enhancements have been made to reduce distortion and power compression, and to optimize airflow in the chambers, producing several unique tweeter designs.

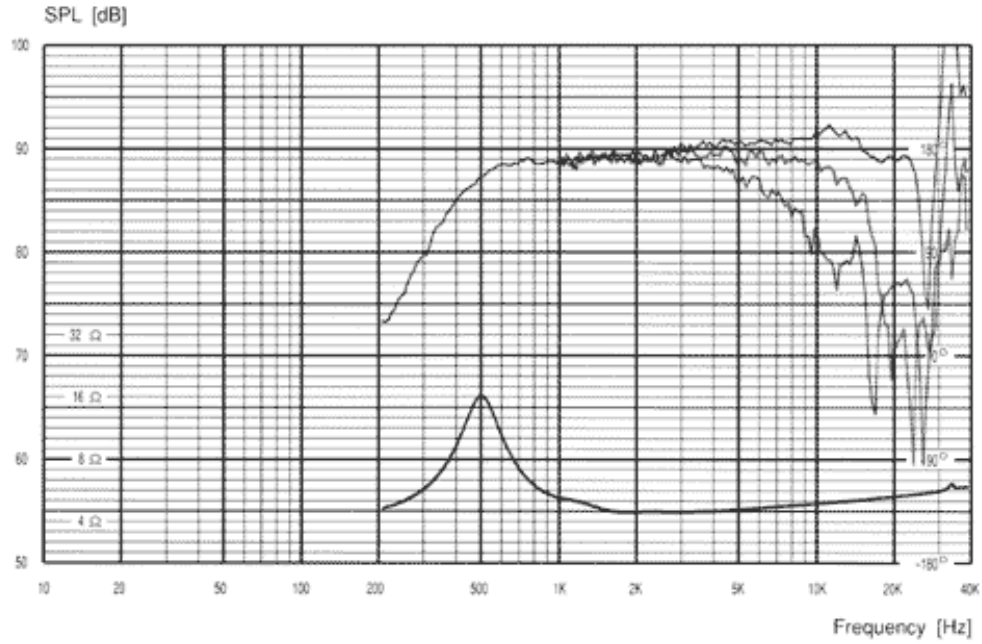
Driver Highlights: 1" aluminium dome, SD-2 motor, low compression design, non resonant chamber

Specs:

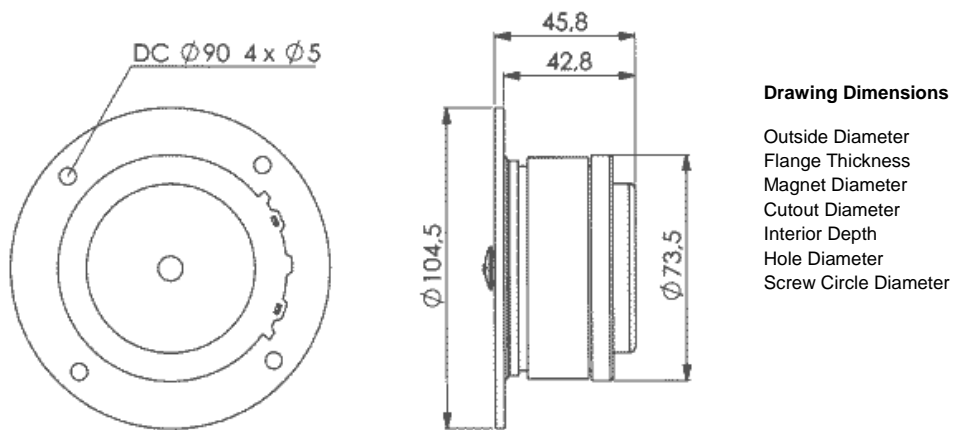
Electrical Data				Power handling		
Nominal impedance	Zn	4	ohm	100h RMS noise test (IEC)	160	W
Minimum impedance	Zmin	--	ohm	Long-term Max Power (IEC 18.3)	--	W
Maximum impedance	Zo	--	ohm	Short Term Max power (IEC 18.2)	--	W
DC resistance	Re	3.5	ohm	Voice Coil and Magnet Parameters		
Voice coil inductance	Le	0.01	mH	Voice coil diameter	28	mm
T-S Parameters				Voice coil height	--	mm
Resonance Frequency	fs	500	Hz	Voice coil layers	--	
Mechanical Q factor	Qms	--		Height of the gap	--	mm
Electrical Q factor	Qes	--		Linear excursion +/-	0.1	mm
Total Q factor	Qts	--		Max mech. excursion +/-	1.5	mm
Force factor	Bl	2.8	Tm	Flux density of gap	--	mWb
Mechanical resistance	Rms	--	Kg/s	Total useful flux	--	mWb
Moving mass	Mms	0.5	g	Diameter of magnet	--	mm
Suspension compliance	Cms	--	mm/N	Height of magnet	--	mm
Effective cone diameter	D	--	cm	Weight of magnet	--	Kg
Effective piston area	Sd	8.5	cm ²			
Equivalent volume	Vas	--	ltrs			
Sensitivity (2.83V/1m)		90	dB			

Notes:
IEC specs refer to IEC 60268-5 third edition.
All ScanSpeak products are RoHS compliant.

Frequency: D2904/980000



Mechanical Dimensions: D2904/980000



M•1400/M•1400i

■ The Mackie M•1400/M•1400i is a dual channel, high output power amplifier that incorporates a number of unique features including FR “Fast Recovery” design, high continuous-current output, built-in subwoofer crossovers, and a T-Design constant-gradient cooling tunnel for improved cooling efficiency and output device reliability.

■ The M•1400/M•1400i is rated at a continuous output of 250 watts per channel into 8Ω, 425 watts per channel into 4Ω and 630 watts into 2Ω. In bridge mode the M•1400/M•1400i is rated at 850 watts into 8Ω and 1260 watts into 4Ω. Variable low-cut filters on each channel, with a range from off to 170Hz, enable tighter bass response and a Constant Directivity Horn EQ improves high-frequency reproduction. The CD Horn EQ is 6 dB/octave and its “knee” position is sweepable from 2kHz to 6kHz. The built-in limiter is channel independent and helps eliminate clipping. Inputs are balanced/unbalanced 1/4" and XLR, and XLR thru connectors allow the full-range signal to be routed to another M•1400/M•1400i. Speaker connections on the M•1400 are Speakon® and binding posts, while speaker connections on the M•1400i are 1/4" and binding posts.

■ To effectively deal with clipping, an amplifier must be able to recover almost instantaneously. That is the definition of “Fast Recovery”. Rather than using negative feedback to help control clipping distortion, the M•1400/M•1400i employs a very sparse amount of negative feedback. The use of Baker Clamp circuits on the positive and negative voltage amp stages prevents saturation (and latching) during periods of overdrive. In addition, a transistor senses when the Baker Clamp is active and activates the internal limiting circuits. This results in no latching, instant recovery from overdriving the amp, and a superior sound.

(continued on page 8)

High-Current Power Amplifier



M•1400/M•1400i

Features

- 1400 watts max, 1260 watts continuous @ 4 ohms bridged
- Ultra-low-noise/low-distortion design
- Fast Recovery circuitry reduces distortion at clipping
- Two 2nd order, 12 dB/octave, Bessel low-cut filters with variable frequency from Off to 170Hz
- Constant directivity horn EQ/Air EQ with variable frequency from 2kHz to 6kHz and on/off switch
- Limiter with on/off switch
- Balanced/unbalanced 1/4" and XLR inputs
- XLR thru outputs
- 1/4" jacks (M•1400i) or Speakon® (M•1400) connectors and 5-way binding post outputs (both models)
- Detented gain controls calibrated in dB and volts
- Signal present and OL LEDs
- Channel Status LEDs
- Superior T-Design cooling
- Five-year, limited warranty (U.S. only)

RELATED PRODUCTS

M•800 Power Amplifier, 1202-VLZ PRO 12-Channel Mic/Line Mixer, 1402-VLZ PRO 14-Channel Mic/Line Mixer, 1604 VLZ PRO 16-Channel Mic/Line Mixer, 1642-VLZ PRO 16-Channel Mic/Line Mixer, SR24•4-VLZ PRO 24 Channel Mic/Line Mixer, SR32•4-VLZ PRO 32 Channel Mic/Line Mixer, CFX Series 12, 16, and 20 Channel Mic/Line Mixer w/Digital Effects, C300z 2-Way Loudspeaker, C200 2-Way Loudspeaker, SWA1501 and SWA1801 Active Subwoofers.

Applications

- Live sound/music reinforcement for churches, clubs, schools, conference centers, hotels
- High level A/V playback
- Large speech systems

M•1400/M•1400i High-Current Power Amplifier

M•1400/1400i Specifications

Maximum Power at 1% THD:

300 watts per channel into 8 ohms
 500 watts per channel into 4 ohms
 700 watts per channel into 2 ohms
 1000 watts into 8 ohms bridged
 1400 watts into 4 ohms bridged

Continuous Sine Wave Average Output Power, both channels driven:

250 watts per channel into 8 ohms from 20Hz to 20kHz, with no more than 0.012% THD
 425 watts per channel into 4 ohms from 20Hz to 20kHz, with no more than 0.025% THD
 630 watts per channel into 2 ohms from 20Hz to 20kHz, with no more than 0.050% THD

Bridged Mono Operation:

850 watts into 8 ohms from 20Hz to 20kHz, with no more than 0.025% THD
 1260 watts into 4 ohms from 20Hz to 20kHz, with no more than 0.050% THD

Power Bandwidth:

20Hz to 70kHz (+0, -3 dB)

Frequency Response:

20Hz to 40kHz (+0, -1 dB)
 10Hz to 70kHz (+0, -3 dB)

Distortion:

THD, SMPTE IMD, TIM
 <0.025% @ 8Ω
 <0.050% @ 4Ω
 <0.150% @ 2Ω

Signal to Noise Ratio

>107 dB below rated power into 4 ohms

Channel Separation:

>80 dB @ 1kHz

Damping Factor:

>350 from 0 to 400Hz

Input Impedance:

20kΩ balanced bridging

Input Sensitivity:

1.23 (+4 dBu) for rated power into 4 ohms

Voltage Gain:

30.25 dB (32.5V/V)

Maximum Input Level:

9.75 volts (+22 dBu)

Rise Time:

<4.4μs

Slew Rate:

Voltage Slew Rate >50V/μs
 >100V/μs bridged
 Current Slew Rate >32A/μs at 2Ω

CMRR:

>40 dB, 20Hz to 20kHz

Transient Recovery:

1μs for 20 dB overdrive at 1kHz

High Frequency Overload and Latching:

No latch up to any frequency or level

Variable Low-Cut Filter:

10Hz (Off) to 170Hz, 2nd Order Bessel

Subwoofer Low-Pass Filter:

Switched: 63Hz/125Hz, 3rd Order Bessel

Constant Directivity High-Frequency Boost:

2kHz to 6kHz (+3 dB points)
 6 dB/octave high-frequency shelving filter, (shelving occurs at approximately 30kHz)

Limiter Section:

Complementary Positive and Negative Peak Detecting

Indicators:

6 meter LEDs per channel
 SIG (Signal Present), -20, -9, -6, -3, OL (Overload)
 PROTECT LEDs, CH1 & 2
 SHORT LEDs, CH1 & 2
 TEMP STATUS
 COLD/HOT LEDs

Power Consumption:

65 watts (0.9A)	at idle
550 watts (6.7A)	with musical program fully loaded (4 ohms per side, or 8 ohms bridged)
900 watts (10.5A)	with musical program fully loaded (2 ohms per side, or 4 ohms bridged)
850 watts (9.6A)	at full power into 8 ohms (continuous sine wave)
1500 watts (15.6A)	at full power into 4 ohms (continuous sine wave)
2500 watts (24.8A)	at full power into 2 ohms (continuous sine wave)

(continued on page 3)

M•1400/M•1400i High-Current Power Amplifier

AC Line Power:

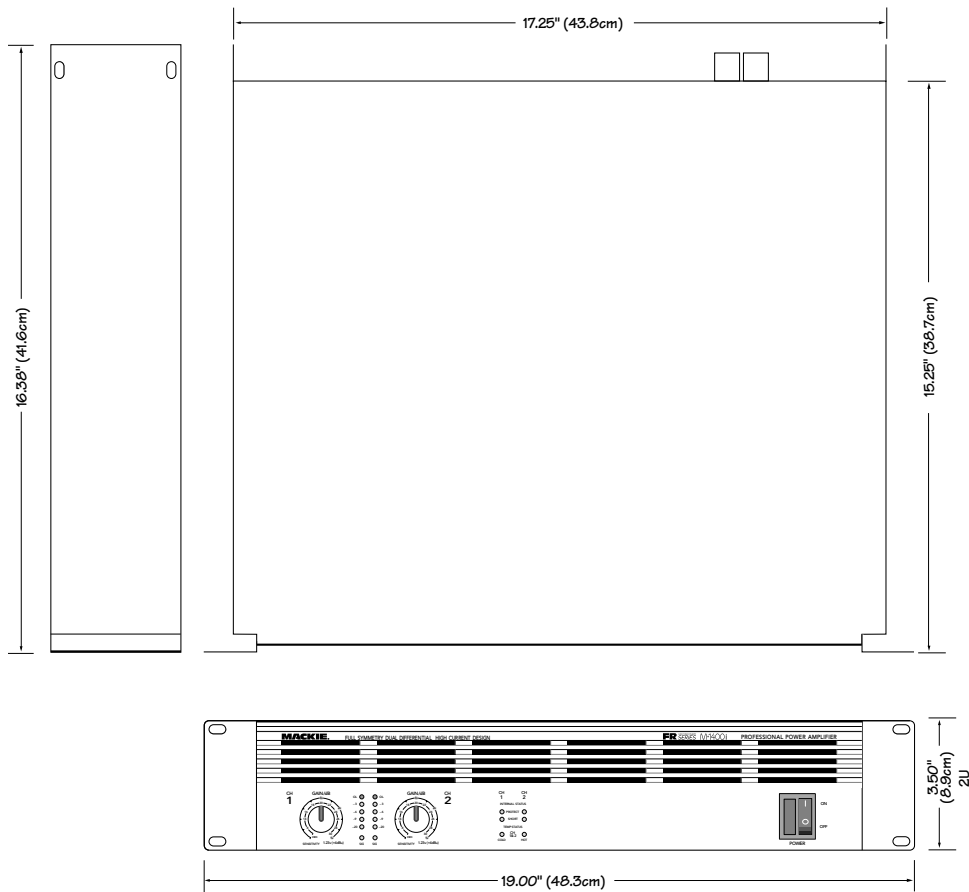
US (M•1400 and M•1400i)	120 VAC, 60Hz
Europe (M•1400 only)	240 VAC, 50/60Hz
Japan (M•1400 only)	100 VAC, 50/60Hz
Korea (M•1400 only)	220 VAC, 60Hz

AC Drop-out Voltage:

At approximately 63% of rated line voltage

Physical:

Height	3.50" (89 mm)
Width	19.00" (483 mm)
Depth	15.25" (387 mm)
Overall Depth	16.38" (416 mm)
Weight	36 lbs (16.3 kg)



MACKIE®

3 OF 8 PAGES


Elektroakustik		 EN 61260	
Bandfilter für Oktaven und Bruchteile von Oktaven (IEC 61260:1995 + A1:2001) Deutsche Fassung EN 61260:1995 + A1:2001			
ICS 17.140.50	Ersatz für DIN EN 61260:1996-02		
<p>Electroacoustics — Octave-band and fractional-octave-band filters (IEC 61260:1995 + A1:2001); German version EN 61260:1995 + A1:2001</p> <p>Electroacoustique — Filtres de bande d'octave et de bande d'une fraction d'octave (CEI 61260:1995 + A1:2001); Version allemande EN 61260:1995 + A1:2001</p>			
Die Europäische Norm EN 61260:1995 zusammen mit der eingearbeiteten Änderung A1:2001 hat den Status einer Deutschen Norm.			
Nationales Vorwort			
<p>Diese Norm beinhaltet die deutsche Fassung der Europäischen Norm EN 61260:1995 und ihrer Änderung A1:2001. Die dieser Europäischen Norm zugrunde liegende Internationale Norm IEC 61260:1995 und ihre Änderung A1:2001 sind im Technischen Komitee 29 „Electroacoustics“ der IEC unter deutscher Mitarbeit erstellt worden. Für die deutsche Mitarbeit ist der Gemeinschaftsausschuss NALS/DKE A 3 „Schallmessgeräte“ unter Federführung des NALS verantwortlich.</p>			
<p>In der vorliegenden Norm ist die Änderung IEC 61260:1995/A1:2001 eingearbeitet und mit einem Strich am Rand gekennzeichnet. Zu dieser Änderung war der deutsche Norm-Entwurf DIN IEC 29/449/CDV:2000-06 veröffentlicht worden.</p>			
<p>Zu den im Abschnitt 2 und im Anhang ZA zitierten Internationalen Normen wird im Folgenden auf die entsprechenden Deutschen Normen hingewiesen:</p>			
IEC 60050(801)	siehe DIN 1320	IEC 61000-6-2	siehe DIN EN 61000-6-2
IEC 60651	siehe DIN EN 60651	IEC 61000-6-3	siehe DIN EN 61000-6-3
IEC 60804	siehe DIN EN 60804	CISPR 16-1	siehe E DIN VDE 0876-16-1/A2
IEC 61000-4-2	siehe DIN EN 61000-4-2	CISPR 22	siehe DIN EN 55022
IEC 61000-4-3	siehe DIN EN 61000-4-3	ISO 266	siehe DIN EN ISO 266
IEC 61000-6-1	siehe DIN EN 61000-6-1		
<p>Diese Deutschen Normen sind in Anhang NA aufgeführt.</p>			
Fortsetzung Seiten 2 und 3 und 34 Seiten EN			
<p>Normenausschuss Akustik, Lärminderung und Schwingungstechnik (NALS) im DIN und VDI DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik im DIN und VDE</p>			

Tabelle A.1 — Bandmittenfrequenzen für Oktav- und Terzfilter im Hörbereich

Index x	Basis zehn: exakte f_m $10^{x/10} \cdot 1000$ Hz	Basis zwei: exakte f_m $2^{x/3} \cdot 1000$ Hz	Nenn-Band- mittenfrequenz Hz	Terz	Oktave
-16 -15 -14	25,119 31,623 39,811	24,803 31,250 + 39,373	25 31,5 40	• • •	•
-13 -12 -11	50,119 63,096 79,433	49,606 62,500 + 78,745	50 63 80	• • •	•
-10 -9 -8	100,00 + 125,89 158,49	99,213 125,00 + 157,49	100 125 160	• • •	•
-7 -6 -5	199,53 251,19 316,23	198,43 250,00 + 314,98	200 250 315	• • •	•
-4 -3 -2	398,11 501,19 630,96	396,85 500,00 + 629,96	400 500 630	• • •	•
-1 0 1	794,33 1 000,0 + 1 258,9	793,70 1 000,0 + 1 259,9	800 1 000 1 250	• • •	•
2 3 4	1 584,9 1 995,3 2 511,9	1 587,4 2 000,0 + 2 519,8	1 600 2 000 2 500	• • •	•
5 6 7	3 162,3 3 981,1 5 011,9	3 174,8 4 000,0 + 5 039,7	3 150 4 000 5 000	• • •	•
8 9 10	6 309,6 7 943,3 10 000 +	6 349,6 8 000,0 + 10 079	6 300 8 000 10 000	• • •	•
11 12 13	12 589 15 849 19 953	12 699 16 000 + 20 159	12 500 16 000 20 000	• • •	•

ANMERKUNG 1 Die exakten Bandmittenfrequenzen wurden nach Gleichung (3) auf fünf signifikante Stellen berechnet mit Ausnahme der mit + gekennzeichneten exakten Werte.

ANMERKUNG 2 In ISO 266 sind weitere Nenn-Bandmittenfrequenzen für Oktav- und Terzfilter angegeben.

Appendix B

Source Code

Appendix C

Raw Data from Subjective Trials

Glossary

acum

unit of sharpness 27

asper

unit of roughness 28

bark

Unit for critical band rate 26

Critical Band

26, 48

dB

decibel 27

FFT

Fast Fourier Transform 20

HMI

Human Machine Interface 12

HUD

Head Up Display 37

JND

Just Noticeable Difference 31

mel

melody, unit of pitch 25

PCB

Printed Circuit Board 29

PSE

Point of Subjective Equality 36

Reverberation time

Reverberation time 47

sone

Unit of loudness 25

Sound Pressure

Sound Pressure 17

SPL

Sound Pressure Level 17

vacil

unit of fluctuation strength 27

Weighting Filter

Filter which models the frequency dependent sensitivity of the human sense of hearing, different filters for different sound pressure levels 18

Bibliography

- [1] D.A. Norman, *The design of everyday things*, Basic Books New York, 2002.
- [2] L. Eckstein, “Automotive HMI: Current Status and future Challenges”, in *Proceedings of AIDE Final Workshop*, 2008.
- [3] J. Reisinger, H. Bubb, J. Wild, and G. Mauter, “Haptical Feeling of Rotary Switches”, in *Proc. EuroHaptics*, 2006, vol. 6, pp. 49–55.
- [4] B. Adelstein, M. Anderson, D.R. Begault, and B. McClain, “Thresholds for Auditory-Tactile Asynchrony”, 2005.
- [5] K. Hotting and B. Roder, “Hearing cheats touch, but less in congenitally blind than in sighted individuals”, *Psychological Science*, vol. 15, no. 1, pp. 60–64, 2004.
- [6] M. Kurosu and K. Kashimura, “Apparent usability vs. inherent usability: experimental analysis on the determinants of the apparent usability”, in *Conference on Human Factors in Computing Systems*. ACM New York, NY, USA, 1995, pp. 292–293.
- [7] N. Tractinsky, “Aesthetics and apparent usability: empirically assessing cultural and methodological issues”, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM New York, NY, USA, 1997, pp. 115–122.
- [8] D.A. Norman, *Emotional design: Why we love (or hate) everyday things*, Basic Civitas Books, 2004.
- [9] P. Glanzmann, “Autopsychologie”, Sat 1 Automagazin 31.03.2007, TV Show.
- [10] H. Fastl and E. Zwicker, *Psychoacoustics: facts and models*, Springer-Verlag New York Inc, 2006.
- [11] C. Patsouras, J. Biber, and W. Ruth, “Gestaltung des subjektiven V8-Sounds”, *Fortschritte der Akustik - DADA 2006*, vol. 32, no. 2, pp. 501, 2006.
- [12] H. Levitt, “Transformed up-down methods in psychoacoustics”, *Journal of the Acoustical Society of America*, vol. 49, no. 2, pp. 467–477, 1971.
- [13] S MacAdams, “Audition: Cognitive psychology of music”, *In The Mind-Brain Continuum*, 1996.
- [14] J.M. Grey, “Multidimensional perceptual scaling of musical timbres”, *Journal of the Acoustical Society of America*, vol. 61, no. 5, pp. 1270–1277, 1977.
- [15] Robert A. Lutfi, Eunmi Oh, Eileen Storm, and Joshua M. Alexander, “Classification and identification of recorded and synthesized impact sounds by practiced listeners,

- musicians, and nonmusicians”, *The Journal of the Acoustical Society of America*, vol. 118, no. 1, pp. 393–404, 2005.
- [16] W.W. Gaver, “Synthesizing auditory icons”, in *Proceedings of the INTERACT’93 and CHI’93 conference on Human factors in computing systems*. ACM New York, NY, USA, 1993, pp. 228–235.
- [17] Daniel J. Freed, “Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events”, *The Journal of the Acoustical Society of America*, vol. 87, no. 1, pp. 311–322, 1990.
- [18] G.J. Hunter, K. Zienowicz, and A.I. Shihab, “The Use of Mel Cepstral Coefficients and Markov Models for the Automatic Identification, Classification and Sequence Modelling of Salient Sound Events Occurring During Tennis Matches”, *Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 3431–3431, 2008.
- [19] D. Ali, “Identification of transient events from a hard disk drive using non-stationary loudness”, *Acoustical Society of America Journal*, vol. 123, pp. 3160, 2008.
- [20] R. Liebing and R. Weber, “Objektivierung des Türgeräusches eines Kraftfahrzeuges”, in *Fortschritte der Akustik - DAGA 2008*, 2008, pp. 109–110.
- [21] E. Altinsoy, “Perceptual aspects of audio-tactile asynchrony”, in *10th International Congress on Sound and Vibration*, 2003.
- [22] D.R. Begault, M.R. Anderson, and B.U. McClain, “Spatially-Modulated Auditory Alerts”, in *Proceedings of the International Conference on Auditory Display ICAD*. Citeseer, 2003, pp. 6–9.
- [23] S. Töpken, M. Bellmann, and R. Weber, “Cross-Modality Matching für Schall und Ganzkörpervibration”, in *Fortschritte der Akustik - DAGA 2008*, 2008, vol. 8.
- [24] J. Chalupper and H. Fastl, “Dynamic loudness model (DLM) for normal and hearing-impaired listeners”, *Acta Acustica united with Acustica*, vol. 88, no. 3, pp. 378–386, 2002.
- [25] D. Menzel, E. Faccinelli, and H. Fastl, “Untersuchung von Farbeinflüssen auf die Lautheit mit einem Einregelverfahren”, in *Fortschritte der Akustik - DAGA 2008*, 2008, vol. 8.
- [26] H. Fastl, T. Rader, G. Van Den Boogaart, S. Kuwano, and S. Namba, “Crosscultural comparison of colour evaluation using semantic differential”, in *Fortschritte der Akustik - DAGA 2008*, 2008.
- [27] A. Amditis, L. Andreone, A. Polychronopoulos, and J. Engstrom, “Design and development of an adaptive integrated driver-vehicle interface: overview of the AIDE project”, in *Proceedings of the IFAC conference, Prague*, 2005.
- [28] J. Huber, “Akustischer simulator”, Term Work, Heilbronn University, 2007.
- [29] A. S. Treiber and G. Gruhler, “Measurement and optimization of acoustic feedback of control elements in cars”, in *Proceedings of the 122nd AES Convention*, 2007.
- [30] A.S. Treiber and G. Gruhler, “Simulation of Room Acoustics Design of a Listening Environment for Sound Quality Jury Tests”, in *Proceedings of Radioelektronika 2006*, 2006.

- [31] C. Roads and J. Strawn, *The computer music tutorial*, MIT press, 2000.
- [32] Eberhard Zwicker, “Dependence of post-masking on masker duration and its relation to temporal effects in loudness”, *The Journal of the Acoustical Society of America*, vol. 75, no. 1, pp. 219–223, 1984.
- [33] Eberhard Zwicker, “Procedure for calculating loudness of temporally variable sounds”, *The Journal of the Acoustical Society of America*, vol. 62, no. 3, pp. 675–682, 1977.
- [34] Eberhard Zwicker, “Temporal effects in simultaneous masking by white-noise bursts”, *The Journal of the Acoustical Society of America*, vol. 37, no. 4, pp. 653–663, 1965.
- [35] Chaoying Zhang and Fan-Gang Zeng, “Loudness of dynamic stimuli in acoustic and electric hearing”, *The Journal of the Acoustical Society of America*, vol. 102, no. 5, pp. 2925–2934, 1997.
- [36] Chaoying Zhang and Fan-Gang Zeng, “Loudness of dynamic stimuli in cochlear-impaired listeners”, *The Journal of the Acoustical Society of America*, vol. 103, no. 5, pp. 3051–3051, 1998.
- [37] Yula C. Serpanos, Honor O’Malley, and Judith S. Gravel, “Cross-modality matching and the loudness growth function for click stimuli”, *The Journal of the Acoustical Society of America*, vol. 103, no. 2, pp. 1022–1032, 1998.
- [38] Robert A. Lutfi, Ching-Ju Liu, and Christophe Stoelinga, “Level dominance in sound source identification”, *The Journal of the Acoustical Society of America*, vol. 124, no. 6, pp. 3784–3792, 2008.
- [39] Robert A. Lutfi, “Auditory detection of hollowness”, *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 1010–1019, 2001.
- [40] John M. Grey and John W. Gordon, “Perceptual effects of spectral modifications on musical timbres”, *The Journal of the Acoustical Society of America*, vol. 63, no. 5, pp. 1493–1500, 1978.
- [41] L.E. Kinsler, A.R. Frey, A.B. Coppens, and J.V. Sanders, “Fundamentals of acoustics”, 1999.
- [42] H. Kuttruff, *Room acoustics*, Taylor & Francis, 2000.
- [43] F. Anwander, “Synthesizer-So funktioniert elektronische Klangerzeugung”, *Aufl. Bergkirchen: PPV Presse Project Verlags GmbH*, 2000.
- [44] H. Kuttruff, “Akustik, Eine Einführung”, *Stuttgart: Hirzel*, 2004.
- [45] A. Nisbett, *The sound studio*, Focal Press, 2003.
- [46] N. Rades, “Bau einer mobilen akustischen prüfkammer”, Term Work, Heilbronn University, 2006.
- [47] N. Otto, S. Amman, C. Eaton, and S. Lake, “Guidelines for jury evaluations of automotive sounds”, *Sound and Vibration*, vol. 35, no. 4, pp. 24–47, 2001.
- [48] E. Hoggan, T. Kaaresoja, P. Laitinen, and S. Brewster, “Crossmodal congruence: the look, feel and sound of touchscreen widgets”, in *Proceedings of the 10th international conference on Multimodal interfaces*. ACM New York, NY, USA, 2008, pp. 157–164.

- [49] M.E. Altinsoy and S. Merchel, “Audiotactile Feedback Design for Touch Screens”, in *Proceedings of the 4th International Conference on Haptic and Audio Interaction Design*. Springer, 2009, p. 144.
- [50] M.E. Altinsoy and S. Merchel, “Audiotactile Feedback for Touch Screens”, in *NAG/DAGA 2009*, 2009.
- [51] J. Blauert and R. Guski, “Critique of Pure Psychoacoustics”, in *NAG/DAGA 2009*, 2009.
- [52] J. Blauert and J.S. Allen, *Spatial hearing: the psychophysics of human sound localization*, The MIT Press, 1997.
- [53] J. Blauert, *Communication acoustics*, Springer, 2005.
- [54] Dr. Ing. h.c. F. Porsche AG, “Porsche cayenne catalog 2007”, 2007.

