International Workshop on Acoustic Echo and Noise Control (IWAENC2003), Sept. 2003, Kyoto, Japan

THE AUTOMATIC DJ

AN APPEALING AND INSTRUCTIVE SIGNAL PROCESSING EDUCATION PROJECT

Piet Sommen, Harrie van Meer, Leo Vogten and Jean Ritzerfeld

Eindhoven University of Technology, the Netherlands.
P.C.W.Sommen@tue.nl

ABSTRACT

Since music and discotheques are appealing to young people, an education project has been defined around the theme "The Automatic DJ". The project groups have to design a programme on their notebook with which they are able to make a smooth transition from one music song, with a certain beat, to another music song, possibly with a different beat, without degrading the quality of the music. The main signal processing challenge is to find and implement an efficient algorithm that performs time scaling (change the beat) without influencing the frequency scaling. The instructiveness of this education project lies in the fact that basic signal processing concepts have to be well understood and applied into a complex real world design problem.

1. INTRODUCTION

Classical courses are important to teach the individual student mathematical and basic concepts. Laboratories can fill the gap between these concept and their physical interpretation. These labs are mainly developed for small groups, upto 2-3 students. An engineer rarely works alone to solve complex real world problems. Education projects are a way to handle such complex design problems in the education process. Students typically work together in groups of 6 - 8 students. These design problems are open ended and allow for a wide variety of possible solutions involving engineering trade-offs between performance and cost. This makes the build up knowledge in courses and labs indispensable: basic courses, labs and education projects form a vital link in the engineering education process. A major idea of education projects is to introduce applications early, well before students have built up enough theory to fully analyse the application. This helps to motivate the students to learn the theory. This is one of the main reasons that since september 2001 the Electrical Engineering (EE) Department of Eindhoven University of Technology (TU/e) has introduced the concept of education projects in the first three years of the curriculum. In each year three different projects are planned in parallel with the usual courses. Each of the nine projects is related to a different basic course and covers a period of 11 weeks.

In this paper we describe the fifth project, which is related to the basic Signal Processing course. Since music and discotheques are appealing to young people, a project is defined in which each project team has to design an "Automatic DJ" programme on their notebook [1]. This programme should create in real-time a smooth transition from one music song, with a certain beat, to another music song, possibly with a different beat, without degrading the quality of the music. Changing the beat, or scaling the time domain samples of the audio signal, inherently influences the frequency scale and thus the frequency content, like pitch, spectral shape, envelope, timbre etc. For this reason the main signal processing challenge of this education project is to find an efficient algorithm that performs time scaling (change the beat) without influencing the frequency scaling. Although such an education project contains many different technical and non-technical aspects we will mainly focuss in this paper on the signal processing (related) topics. The other issues are only mentioned and described very briefly.

This paper is organized as follows: First the main objective of the project is described in section 2, while section 3 gives a global overview of the project organization, the global schedule and the support. Lab exercises are important to realize the project. Different technical and communication skills are discussed in section 4, while section 5 gives a description of the signal processing labs. The construction of the individual final assessment is discussed in section 6 and the conclusions are given in section 7.

2. MAIN OBJECTIVE

Education projects brings synergy in the tight coupling of project and course material, i.e. project experience reinforces lecture theory and techniques learned in lectures often motivates project strategies. For this reason each project runs over a period of 11 weeks in parallel to the regular courses. The study load of he project is approximately 30–35% ($\approx$ 180 hours) and each project

[1] At TU/e all students own a notebook (@1.4 GHz in 2002) or have one at their disposal

is related to a specific basic course. Furthermore an engineer rarely works alone. Learning to collaborate effectively is an important part of the working world. In view of this the students typically work together in groups of 6 – 8 students, with each group attacking the same design problem. The design problems are open ended and allow for a wide variety of possible solutions involving engineering trade-offs between performance and cost.

The main objective of this signal processing related project is that at the end of the project each student should be able to:

- Analyse and physically understand some important basic signal processing concepts,
- Execute an audio signal processing algorithm in real-time,
- Design a complex software programme within a group of students,
- Work together in a team,
- Design a realizable solution for a complex signal processing problem,
- Report and communicate the main project results by means of writing a paper, making a poster and give a presentation/ demonstration.

3. THE "AUTOMATIC DJ" PROJECT

Each (audio) signal can be described in both time- and frequency-domain. A basic property of the scaling in one domain is the inverse scaling in the other domain. This is represented in Figure 1. The upper-left figure represents the original audio signal while the lower-right figure is a scaled version of it [2]. One of the main signal processing challenges of the "Automatic DJ" project is to perform to a given audio signal an efficient time scaling (beat control) in real-time, while not scaling the frequency content. This situation is represented in the upper-right corner of the figure. Finally the lower-left corner of the figure represents a scaling of the frequency (pitch control), without scaling the time axis. Although an interesting option, e.g. to change the pitch of a speech signal, pitch control is not included in the project. The "Automatic DJ" assignment is globally defined as follows:
Design with your team an "Automatic DJ" programme on your notebook that performs an automatic scaling from one music song to the other. The programme should be able to make in real-time a smooth transition from one music song with a certain beat to another music song, with a possible different beat, without degrading the quality of the music.

[2]In order to represent the time-frequency relation in an exact way we could have used a music staff. Although not exact, we found it more appropriate and representative to use a part of the IWAENC logo with 'virtual' time- and frequency axis.

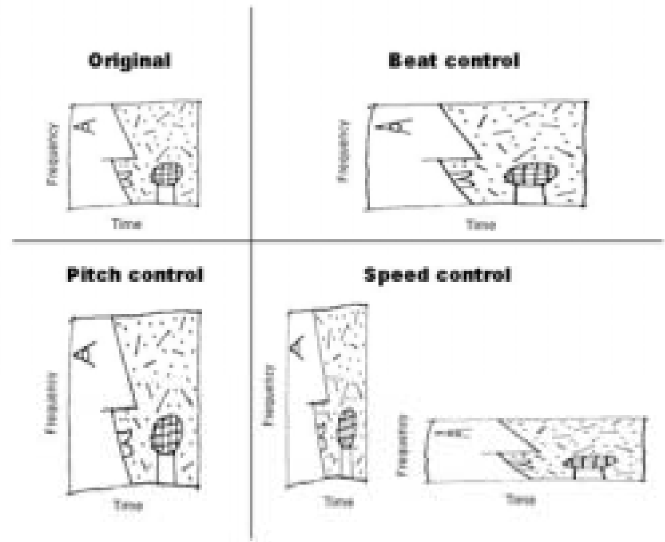

Figure 1: *Four different time- frequency scaling situations*

3.1. Global schedule

A global schedule of the project is given in Table 1. In the introduction meeting (Intro) the "Automatic DJ" project is introduced and roughly described. The groups (6 – 8 students) are composed and each group is assigned their own tutor and their own office to carry out the project. Within one week each group selects its own student project leader. The project is roughly divided into three phases (see Table 1): The preparation phase, the design phase and the concluding phase.

The preparation phase, ≈ 35 % of the study load of the whole project, is oriented bottom-up and the main goal for each individual student in the group is to acquire basic signal processing knowledge and to obtain technical and communication skills. For this reason five signal processing labs (SP1 – SP5), three technical skill labs (TS1 – TS3) and three communication skill labs (CS1 – CS3) must be executed by all students in the preparation phase of the project. The skill labs, described into more detail in the following section, are scheduled at the moment they are needed. E.g. an introduction to the use of Matlab is treated before the first signal processing lab. The preparation phase is concluded with a first short individual oral exam (IE1).

The design phase, ≈ 45 % of the study load of the whole project, is oriented top-down: each group makes its own design, from concept to realization and test. For this reason each group has to write and defend a plan of action in the first project evaluation meeting (PE1). Besides a final result this plan must also describe an intermediate result, that is evaluated in

Preparation phase (*Load* ≈ *35 %*)	
Intro	Intro project, composition groups
TS1	Intro Matlab
SP1	Basic signal processing
CS1	Team building 1
SP2	Spectral Analysis
TS2	Audio Perception
CS2	Group leader training
SP3	Beat Detection
SP4	Speed Control
SP5	Beat Control
CS3	Team building 2
TS3	Real Time Audio Processing
IE1	Individual exam 1
Design phase (*Load* ≈ *45 %*)	
PE1	Plan of action
PE2	Intermediate result
IE2	Individual exam 2
Concluding phase (*Load* ≈ *20 %*)	
PE3	Paper
PE4	Presentation
PE5	Demonstration
PE6	Poster
PE7	Software

Table 1: *Scheduling and study load of the three phases in the "Automatic DJ" education project and their subdivision into technical skill labs (TS), signal processing labs (SP), communication skill labs (CS), individual exams (IE) and project evaluation steps (PE).*

project evaluation (PE2). The design phase is concluded with a second short individual oral exam (IE2).

The concluding phase, ≈ 20 % of the study load of the whole project, is conference-like: each group has to write a short 4-6 page paper (PE3), give a short presentation (PE4) with a demonstration of its final product (PE5) and present a poster (PE6) in a poster session. The sources of the software are judged separately (PE7).

In Figure 2 a setup is shown of the equipment that is used to realize the "Automatic DJ" education project. The developed algorithm runs in real-time on the notebook using music songs that where available on it. The output is transferred via standard audio devices, a high quality (44.1 kHz, 16 bit) digital USB/SPDIF interface, to a standard set of speakers.

3.2. Support

The organization issues of the whole project are handled by the project coordination consisting of a project coordinator and a project administrator. The support

Figure 2: *Photo of the "Automatic DJ" setup*

and guidance of the groups is divided into a technical guidance part and a group-process guidance part. The technical guidance part is mainly taken care of by a so called Team of EXperts (TEX-team), which consists of 2 – 3 persons: the lecturer of the basic signal processing course, the project coordinator and an expert in the subject under study at that specific moment. Furthermore a helpdesk, parttime manned by PhD students, is available during project hours for direct and quick technical support. The group-process is mainly guided by the tutor and during the design phase once a week there is a group leader meeting under supervision of a communication expert and the project coordinator.

During the whole project each group has on a regular basis once a week a 45 minutes during meeting with the TEX-team. In the preparation phase these meetings are mainly used to discuss the signal processing lab results, while technical aspects of the projects are discussed in the design phase. The project groups are stimulated to use their creativity and/or level of sophistication in their approach. The main advantage of the set-up with a TEX-team is twofold: 1) the project coordination keeps an overview of the technical advances of the project groups 2) the critical discussions during the TEX-team meeting stimulates a positive technical attitude during the project.

The main task for the tutor is to track, and if necessary guide, the group-process by paying attention to the following aspects: is the project leader functioning well; are all members doing their task; is the group process diverging or converging; are there hanger-on students; etc.

The written communication between the project coordination and the project groups takes place by a well structured email folder, that is owned (full access) by the project coordination. Besides a general information folder and an assignment folder each group is assigned an own group folder. This group folder is split into an internal group-data folder and a group-result folder, for writing (no editing/ no deleting) all the assignment results.

4. SKILLS LABS

As shown in Table 1 this project has technical and communication skill labs. Since this paper is mainly signal processing oriented, the communication skill labs (CS1 – CS3) that deal with team building and group leadership, will not be treated into more detail here. The technical skill labs TS1 – TS3 are executed by groups of 2 students each.

4.1. Matlab introduction(TS1)

During the project Matlab is mainly used as an analysis tool while PowerPoint [3] is mainly used as a presentation tool to exchange technical ideas, suggestions and possible solutions. In order to get a quick start, TS1 is organized as a brush-up for the use of Matlab and PowerPoint. A primitive way of changing the tempo of an audio signal, as depicted in Figure 3, is used as a vehicle. In this method blocks of L samples are selected

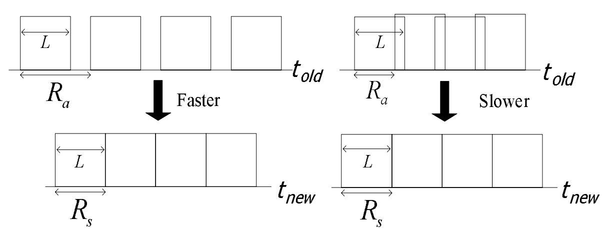

Figure 3: *Processing of blocks*

from the input audio signal. In order not to change audio frequencies in this block, these L samples are sent unchanged to the output. At the input jumps are made of R_a and at the output of R_s samples. In order not to obtain 'holes' or 'strong' parts in the output signal R_s is chosen equal to L. By using the definition $\alpha = R_s/R_a$ and by changing R_a the tempo can be increased when $\alpha < 1$ and decreased when $\alpha > 1$. The assignment of TS1 is as follows:
Implement in Matlab this primitive way of changing the tempo of an audio signal by writing a function that is declared as follows:

```
y = simple(x, alpha, L)
```

in which x represents the input audio samples and y the output samples. Test this function and judge the quality with some predefined (or own) audio files. Each group must prepare one or two PowerPoint slides containing: a) The Matlab code (≤ 15 lines) b) A properly defined plot of an input and an output sequence of 1000 samples and c) A demonstration of the audio quality by using the audio icon of PowerPoint.

[3]Students have used Matlab® and PowerPoint® earlier in the EE curriculum at TU/e.

4.2. Perception (TS2)

The aim of this Technical Skill lab is threefold: 1) To gain basic insight in the most important aspects of audio perception. 2) To get some skills in recognizing the problems and complications that arise when audio quality has to be assessed 'objectively'. 3) To formulate a recommendation for a method to evaluate the music output of the Automatic DJ. Some background can be found in [3], ..., [5].

TS2 consists of two parts: a presentation of the basic principles of audio perception and an overview of the most important aspects of audio quality. The perception part briefly introduces the anatomy of the ear, the principles of cochlear mechanisms that transduce sound into nerve signals, the concept of masking and a simple a model describing this transduction in terms of a bank of band pass filters. The part on audio quality is intended to make the students aware about the difficulties when audio quality should be assessed in a reproducible and generally accepted way. A few objective and subjective standardized methods are presented. Students are invited to think about questions like: which factors determine audio quality, how can audio quality be defined and how can it be measured and how should systems for processing and manipulation of audio signals be evaluated? At the end of the TS2 each group of two students submits a written proposal to assess the audio output of the developed Automatic DJs.

4.3. Real-time audio processing (TS3)

At the end of the design phase each project team must implement an efficient real-time programme for the "Automatic DJ" on their notebook in the programming language C by using a programme development environment.

To reach this purpose students use the programme development environment "The Borland C++ Builder" and the supporting C-Library: "Telecommunications & Signal Processing C-Library" [4](*libtsp*) for the Signal Processing Procedures. Borland® C++Builder^{TM} is a powerful ANSI C++ development environment for rapidly building applications with middleware flexibility and open standards. The *libtsp* package is a library of routines for signal processing. It also includes a number of general purpose routines useful for programme development.

For the real-time implementation use is made of the "Portaudio C-Library for audio input and output" [5]. *PortAudio* is a cross platform, open-source, audio I/O library in which the user can write simple audio programs in C that will compile and run on e.g. Windows. *PortAudio* project and Advanced Programming Inter-

[4]www.tsp.ece.mcgill.ca/MMSP/Documents/Software/index.html
[5]www.portaudio.com

face (API) is intended to promote the exchange of audio synthesis software and sound file support between developers on different platforms. It provides a very simple API for recording and/or playing sound using a simple callback function.

Skill Lab TS3 is executed by groups of 2 students going through the following steps: a) Introduction to Console- and Windows-applications by designing some small programs for reading and writing wave-files. b) Learn to use C-libraries for real-time output resulting in a programme for playing wave-files. c) Design a Windows Application in C and both link the *PortAudio* and *libtsp*. The example used is to filter an audio wave-file with a simple notch filter, from *libtsp*, and to play the result in real-time on an audio output device. d) Check the processor load during all the steps and take care about high loads.

5. SIGNAL PROCESSING LABS

What follows from Figure 1 is that there is a need to first acquire some basic signal processing (see also [1, 2]) knowledge in order to be able to design an "Automatic DJ" programma. This goal can be reached via the signal processing laboratories (SP1 – SP5) that are described in this section. These labs are very important to understand the connection between a mathematical and a physical interpretation of algorithms. Because of the analytical nature of these labs, they all have been carried out in Matlab.

To illustrate the type, scope, and complexity of subjects the students are required to carry out in the signal processing labs, brief descriptions of the subjects are provided in the following subsections. In each lab the group leader divides the project group into different subgroups. Each subgroups studies one of the subjects and prepares a few slides describing the mathematics and the physical interpretation. A motivation of the relevance of the studied subject within the project must be given together with a representative audio example. Each subgroup presents and discusses the results with colleague students. All resulting slides are tuned and emailed to the project coordination, before a predefined deadline. These slides are used as a basis for discussion with the TEX-team. The assignments of the first labs are rather detailed, while the later labs have a more open character.

5.1. Basic audio signal processing (SP1)

As mentioned before, students have to use their notebook. All used audio signals have a digital format: 16 bits stereo at 44.1 kHz. Furthermore the understanding of the relation between time- and frequency-domain scaling of an audio signal is very important within this project. Ultimately in order to make the "Automatic DJ" really work, some way of processing of the audio signals has to be done. Filtering (convolution) and (cross)correlation are the most important examples of such processing operations.

SP1 starts with a plenary meeting in which examples of all these basics signal processing are shown and discussed. The assignment SP1 is as follows:
a) Explain to each other, both in a mathematical and physical interpretive way, how to choose the sample rate f_s of a discrete-time audio signal in relation to the highest available frequency of this signal in order not to obtain aliasing. Show this effect by implementing a simple Matlab function and test this function with at least two predefined audio signals. First use a sine wave of 2kHz which is sampled with $f_s = 5kHz$, $f_s = 4kHz$ and $f_s = 3kHz$. The second example is a predefined audio signal (part of a violin concert) which is originally sampled at $f_s = 44.1kHz$ and which should be re-sampled at $44.1/8kHz$.
b) Explain to each other, both in a mathematical and physical interpretive way, how time-scaling is related to frequency-scaling. Show this scaling effect by implementing a simple Matlab function and test this function with at least two predefined audio signals.
c) Explain to each other, both in a mathematical and physical interpretive way, that the convolution of an input signal with an impulse response is equivalent to a filter operation. Show this filtering operation by implementing a short Matlab programme that performs a pre-defined low-pass filter and test this function with at least two predefined audio signals.
d) Explain to each other, both in a mathematical and physical interpretive way, how the convolution and correlation operation are related to each other. Show the correlation operation by implementing and testing a short Matlab programme for at least two cases. With the first one you must be able to detect a known signal in noise. The second one must find the periodicity in a predefined signal.

5.2. Spectral analysis (SP2)

The main goal is to study and understand a) The tradeoff between window length and bandwidth of a chosen window function b) Spectral analysis properties and resolution in both time- and frequency-domain of the short-time Fourier transform.

SP2 starts with a plenary meeting in which a demonstration is given of the influence of length and shape of time windows upon the amplitude spectrum of various signals. In the remaining part of the afternoon each group executes a number of assignments, using Matlab. The assignments consist of writing short Matlab scripts to calculate and plot the amplitude spectra of various time windows, different in length and different in shape (Rectangular, Bartlett, Hann, and Hamming windows). The resulting plots are interpreted, to illustrate the trade off between time and frequency resolution. Then the various windows must be applied on time varying signals. Finally the concept of spectrogram is introduced and must be interpreted for a stationary sine wave, a step-wise varying sine wave, a speech signal and a music signal.

5.3. Beat detection (SP3)

Before we are able to change the tempo of an audio signal, one should first detect the beat. As a human being we are used to detect the beat by simple foot-tapping and counting. Within this project this needs to be done in an automatic way. For this reason all students of the group are first guided through a pre-processing stage in which only the relevant beat information is extracted from the original audio signal. This pre-processing stage can be described roughly as follows: It is obvious that the beat is present in the low frequency part of the audio signal. We are interested in the smooth envelope of this low-pass filtered signal. Perceptively the timing of an audio event is typically at the beginning of each event, the so-called onset. This is also the case with the onset of a beat, which is at the beginning of the largest increase of the smooth envelope of the low-pass filtered signal. The resulting signal represents a rough measurement of the onsets of the original audio signal, for which there exists different ways to detect the beat. The assignment of SP3 is as follows:
Calculate by foot-tapping the beat of the predefined audio signals and use these numbers as a comparison of your final results. Implement in Matlab a function that performs in an efficient way the described pre-processing stage of the audio signal. Test this function and judge the quality for at least two pre-defined audio signals. Use the resulting signal of this pre-processing stage to study at least the following three approaches to detect the beat: a) The autocorrelation function b) A bank of predefined resonators and c) The histogram function by giving each beat one label. Analyse, implement (in Matlab), test and explain the results. Make, by taking into account accuracy and complexity, a choice for the beat detector that you want to use for the "Automatic DJ" project.

5.4. Speed control (SP4)

Since both the input and the obtained output signals are available at 44.1 kHz., speed control can be applied by using up- or down sampling techniques. The main objective of SP4 is that students come to the understanding that the whole range of rational factors for sample-rate conversion can efficiently be implemented using a fixed lowpass filter in between up- and a down-sampler, as depicted in Figure 4. The assignment of SP4 is:

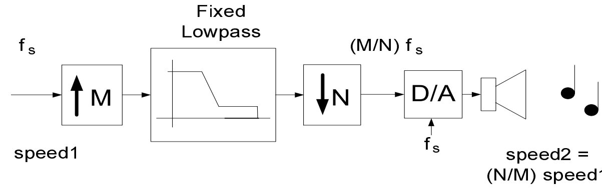

Figure 4: *Speed control with fixed lowpass filter*

Design and test an efficient Matlab programme that reads an audio signal at speed1 and writes at speed2. The speed control should span a range of 50% (up and down). Thus speed2 must be controllable, during run-time, in a range of 2/3 – 3/2 of speed1, or between 67% and 150% of the original speed.

5.5. Beat control (SP5)

The main goal of this lab is to study and understand some different mechanisms to control the beat without changing the frequency content. This is done by using the processing of blocks (see Figure 3), as treated in TS1. For a periodic signal this can be achieved very easily by cutting out or pasting exactly one or more periods of the signal as depicted in Figure 5. If the cut

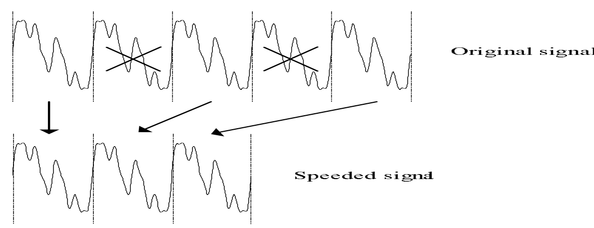

Figure 5: *Simple beat control for a periodic signal*

does not exactly fit to one or more periods one can either shift the signal of the new block, as depicted in Figure 6, or find another new block that fits better. These

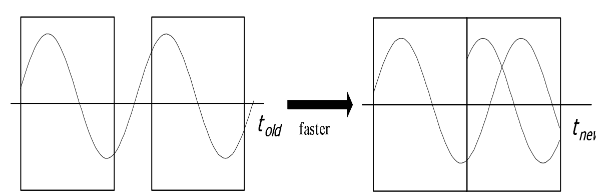

Figure 6: *Shifting the sine wave in the new block to fit the previous sine wave.*

two approaches are the basis of two different classes of beat control algorithms known in literature. The one is the Time Domain Harmonic Scaling (TDHS) approach ([6], ..., [12]) which tries to find an optimal fitting of consecutive segments (blocks) in the audio signal by means of cross-correlation between these consecutive segments (blocks). The other approach is the Phase

Vocoder (PV) ([13], ..., [19]): By applying a filterbank, this method tries to restore the original phase in each frequency band of consecutive segments (blocks) of the time scaled signal. The assignment of SP5 is now as follows:
Use the processing of blocks programme of TS1 as a basis and use two pre-defined audio signals. The first one is a stationary sine wave. The second one represents a slow frequency sweep in such a way that the sweep ranges from 2/3 until 3/2 of a period in a selected block. Design and test two Matlab programmes that change the beat in a pre-defined way without changing the pitch with the following different methods: a) By finding the optimal fitting of consecutive blocks by using the cross correlation function between (parts of) different blocks. b) By shifting the phase of the signal in each new block in such a way that it fits the phase of the signal in the previous block.

6. FINAL ASSESSMENT

Without going into too much detail, this section describes shortly the evaluation of different steps in the project. From this the construction of the final assessment of each individual student is given, as depicted in Table 2. From this table it follows that there are

What	*Weight*	*Jury*
Individual Technical ($\overline{\text{IE}}$)	40%	Proj. Coord./ Tutor
Group Technical ($\overline{\text{PE}}$)	30%	Proj. Coord./ Tutor/ Comm. expert
Group process	20%	Tutor
	10%	Other group members

Table 2: *Construction of individual final results with* $\overline{\text{IE}}$ *and* $\overline{\text{PE}}$ *the average of the different IE's and PE's*

three main categories, namely a) The Individual technical component, which counts for 40% b) The Group technical component, which counts for 30% and c) The Group process component, which counts for 30%.

The first component ($\overline{\text{IE}}$) is the average of the two individual oral exams (IE1 and IE2). The first individual exam IE1 reflects the individual knowledge of all skill and signal processing labs that takes place in the preparation phase (see Table 1). The results of the skill and signal processing labs themselves are not counted in the final assessment. Only participation of these labs is obliged. All results of the labs are discussed and coupled back to the participants. The second individual oral exam reflects the individual knowledge of the project. During both exams the jury is composed of the TEX team members and the tutor of the group to which the student belongs.

The second component is an average of all project steps (PE1 – PE7). It reflects the technical- and communication skills within the group. The judgement is taken care of by the TEX team, the tutor of the group and the communication expert.

The last component, the judgement of he group process, is split into two parts: The tutor gives each group member a final assessment, which counts for 20% and each group member gives the other group member a mark. These marks are averaged and weighted for 10%.

7. CONCLUSIONS

In general students find the " Automatic DJ" project an appealing, difficult and instructive education project.

During the project we have seen a positive attitude change by all participants. A complex realistic real-world problem stimulates students on the one hand to work together in a group, while on the other hand a structured approach has to be followed to tackle such a complex problem.

A point of improvement is the need to learn how to develop a complex software programme with a group. For this reason next year a new Technical Skill lab (TS4) will be developed with this theme.

Furthermore the individual assessment of student working together in a group stays a difficult issue. In this project we solved this (partly) by using enough coaching capacity (TEX-team meetings, tutors, group-leader meetings, individual oral exams etc.), which makes education projects rather expensive: The first time (preparation + run) costed in total $\approx$ 2 fte (full-time-equivalent), while it is expected that in the long run it will cost $\approx$ 1 fte for a group of 50 - 75 students.

Although there are still points of improvement, the main conclusion is that courses, labs and education projects form a vital link in the modern engineering education process.

8. ACKNOWLEDGEMENT

The following persons are acknowledged for their help during the design-, preparation- and running-phase of the "Automatic DJ" project:
Emanuël Habets, Harrie Kuipers, Herman Ossevoort, Wil Hendrix, Connie van den Bergh, Erik Franken, Jasper Schrader, Ronald Aarts, Armin Kohlrausch, Ad van den Enden, Werner Oomen, Frans de Bont, Frank Zijderveld, Jakob van de Laar and Ali Nowbakht. Last but not least, Jeroen Veen is acknowledged for bringing in the original "Automatic DJ" idea.

9. REFERENCES

[1] John G. Proakis and Dimitris G. Manolakis. Digital Signal Processing Principles, Algorithms, and applications, third edition. *Prentice Hall International Editions*, ISBN 0-13-394338-9.

[2] Sanjit K. Mitra. Digital Signal Processing A computer based approach. *McGraw-Hill International Edition, Electrical Engineering Series*, ISBN 0-07-118175-X.

[3] Watkinson, J. The art of digital audio *Focal Press, Oxford*, ISBN 0-240-51587-0; pp38-61: "Some audio principles"; pp704-739: "Sound quality considerations".

[4] Kleijn, W.B., Paliwal, K.K. (1998) Speech coding and synthesis *Elsevier, Amsterdam*, ISBN 0-444-82169-4; pp469-494: "Evaluation of speech coders".

[5] Dimolitsas, S. Subjective assessment methods for the measurement of digital speech coder quality *Speech and Audio Coding for Wireless and Network Applications. Ed. by B. Atal et al., Kluwer Academic Publishers*, 1993, pp. 467–494.

[6] Laroche, J. Autocorrelation Method For High-Quality Time/Pitch-Scaling. *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1993. Final Program and Paper Summaries.
New York: IEEE, 1993. pp. 131–134

[7] Lent, K. An efficient method for pitch shifting digitally sampled sounds. *Computer Music Journal*, Vol. 13 (1989), No. 4, pp. 65–71.

[8] Moulines, E. and J. Laroche Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech Communication*, Vol. 16(1995), pp. 175–205.

[9] Parviainen, O. The PaceMaker plug-in home page. www.sunpoint.net/ oparviai/pacemaker or www.sunpoint.net/~oparviai/pacemaker

[10] Rabiner, L.R. and R.W. Schafer Digital Processing of Speech Signals. *Englewood Cliffs : Prentice-Hall*, 1978. pp. 141–158.

[11] Verhelst, W. and M. Roelands An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech. *Proc. IEEE 1993 International Conference on A coustics, Speech and Signal Processing*, Minneapolis, 27-30 April 1993. New York: IEEE, 1993. Vol 2, pp. 554–557.

[12] Verhelst, W. Overlap-add methods for time-scaling of speech. *Speech Communication*, Vol. 30(2000), pp. 207–221

[13] Dolson, M. The Phase Vocoder: A Tutorial. *Computer Music Journal*, Vol. 10 (1986), No. 4, pp. 14–27.

[14] Flanagan, J.L. and R.M. Golden Phase vocoder. *The Bell System Technical Journal*, November 1966, pp. 1493–1509.

[15] Hammer, F. Time-scale Modification using the Phase Vocoder. An Approach based on deterministic/stochastic component separation in frequency domain. *Institute for Electronic Music and Acoustics*, Graz (Austria), 2001. www.kug.ac.at/iem/lehre/arbeiten/hammer.htm or www.kug.ac.at/iem/lehre/arbeiten/hammer.htm

[16] Laroche, J. and M. Dolson Improved Phase Vocoder Time-Scale Modification of Audio. *IEEE trans. on Acoust., Speech and Signal Processing*, Vol. 7(1999), No. 3, pp. 323–332.

[17] Portnoff, M.R. Implementation of the digital phase vocoder using the fast fourier transform. *IEEE trans. on Acoust., Speech and Signal Processing*, Vol. ASSP-24(1976), pp. 243–248.

[18] Puckette, M. Phase-locked Vocoder. *Proc. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, 15-18 October 1995. New York: IEEE, 1995. pp. 222–225.

[19] Serra, X. and J.O. Smith A system for Sound Analysis/Transformation/Synthesis based on a Deteministic plus Stochastic Decomposition. *Proc. Signal Processing V, Theories and Applications. Proceedings of EUSIPCO European signal processing conference: 5th European Signal Processing Conference*, Barcelona, 18-21 September 1990. Ed. by L. Torres et al. Amsterdam: Elsevier, 1990. Vol. 2, pp. 1347–1350.