Musical Analysis of Computer Music with Sonograms

Martha Brech
Technische Universität Berlin
H 51
Str. des 17. Juni 135
D - 1000 Berlin 12

Abstract: Computer generated sonograms may serve as a base for analysis for all kinds of music, especially computer music, which cannot be notated otherwise. Considerations about the handling of sonograms are discussed in this paper as well as an analysis system for sonogram notation and first results are presented.

Introduction

In this paper some ideas will be presented which arose during a musicological project about experimental computer and electronic music. It deals with understanding and perception of this kind of music, asks for the leading parameters of it and the necessary musical analysis.

For musicologists analysis of experimental electronic and computer music is a problem for musicologists due to the lack of formal notation. Usually, some sort of listening transcription are made to suit the interest only of the individual researcher. Moreno (1985) and Smalley (1986) gave listening and analyzing systems with fixed definitions. These works help a lot, but the matter is still very complicated and contains the problem of all sensual transcription: the lack of objectivity.

Sonograms as Notation

Sonograms may be helpful here. They show:

- on x axis: the passing time
- on y axis: frequencies, the amplitude of each frequency is shown by the printing density

That means that sounds are to be seen as visual patterns which even can be separated from each other in case they overlap. So they can be used as a kind of (postscriptive) notation. The idea of using sonograms as notation is as old as the analogue sonograph. R. Copas even gave an analyzing system based on photographed sonograms at the oscilloscope, but only computer generated sonograms are able to offer sufficient solution with printouts. At the Electronic Studio at the Technical University, Berlin, an FFT-based sonogram program, inspired by Tomas Urayvary and his project ideas (see DOM, 1980), was written on Microvax by Wilger Beckner. It is able to analyze frequencies up to 16 kHz and to print them linearly or logarithmically (at any logarithm base). There are 16 dynamic steps in a variable dynamic range up to 120 dB. The representation of time can vary up to a meter or more per second. (The actual length is due to the variation of FFT-size and overlap.)

It turned out that the most effective printout of a usual sonogram had the following parameter settings: spectrum: in octaves, dynamic range: 60 dB (=3.75 dB per step), time: 13 cm per second.

Reading and Understanding Sonograms

Eye and ear are different perception organs. Although reading of sonograms usually is easy, it must be learnt.

The printouts show all frequencies in a given sound equally, the fundament tone as well as the overtones, so that one sound may be seen as a pattern with multiple ranks. This will not only be a problem, if many different sounds are involved - in that case the printout will turn into a non-readable grey (e.g. Baker II by Xenakis looks like this) - it also means that all partial tones are printed with their absolute amplitudes. Therefore
a sound with relatively fewer overtones is printed darker than a sound with a lot of overtones even if the ear perceives them as equally loud. Additionally, in hearing perception, frequencies are filtered. Some frequencies may occur louder to the eye (because they are printed darker) than they are heard. There may be even cases where a pattern appears to be overwhelming by looking at, but is not heard at all. (The sonogram shows such a sound, a second sound overlap, but the individual analysis cannot be heard, it is masked by lower frequencies.) Filtering the sonogram through the listening perception curve is no real solution because sound patterns, while overlapping, are sometimes easier to be identified in upper frequencies. Since a sonogram is a quantitative notation, listening to the music while reading and analysing seems to be the best approach.

Musical Analysis with Sonograms

As prescriptive notation, sonograms print out the sounding result of a piece of music. For computer music this implies that the sonograms will be able to separate a lot of different sounds but tell nothing about their creation. On first sight, this might be difficult because the main emphasis in computer music lies in the creation of new sounds. In fact, in most books about computer music there is not much more to learn. But putting together these sounds makes the music. The sound itself is only part of a piece. Therefore, from a musicological point of view, the question of the relationship between sound and its construction is of interest too.

In a communication system between composer and audience sonograms are placed on the side of the audience. Hence, the following question arises: What is reaching the audience, what can be perceived. Later it can be asked, if some of the sonograms' findings correlate with the composer's intention.

The Perceiver's View

Perceiving is an individual process which is based both on experience and human capacity in information processing. Experience here means either cultural nor personal experience of life or other aspects of apperception, it just implies that in this field of music there are either expert or untrained listeners. Therefore, the evidence of such an individual analyzing result is questionable. Speed of learning must be taken into account. Stoffler (1985) e.g. investigated the perception of structure (in tonal music) by a click test. He compared two groups, one with high musical competence and training and another with low musical competence. Much to his own surprise the second group seemed to have learned during the test: In contrast to the first session's results Stoffler discovered only little difference between the two groups after four sessions. As analyzing involves frequent listening to a piece, learning can even take place there and hence will be less individual as assessed. Thus, the perception of the senses - ear and brain, the simple biological facts, which can be explored experimentally, are of interest. In most cases, investigations are made with tonal music. However, some are valid for the perception of computer music too, like the fusion of sinus tones to musical tones or duration and remembering principles etc. Prause (1982) writes of a maximum of five perceiving channels depending on the nature of the events. They are limited to 2 or 3 when only duration is concerned. Although rhythm is usually of lesser interest in computer music, duration of sound layers may be important, especially in perceiving structures. 1982 Erickson wrote of a maximum of three melodic streams which can be maintained while listening. In my experience this corresponds to analyzing computer music: Up to now it was impossible to distinguish more than three layers of sound. If new sound qualities
occur, they mix with the ones already heard into a sound layer with the nearest similarity
to the new sound. Thus, even three different layers of sound can melt into one.
Perceiving structures is another important feature. In most cases a structure may be fully
understood only after listening to the piece several times, like Pollard - Gott researched
(1983), if it can be understood fully at all. Here, research results which deal with
remembering are helpful.
While analyzing the music with sonograms it will be senseless to take all features in
account. With the exception of contradictory and unclear parts, a piece can be analyzed
as fundamentally as possible and later one can ask, if all parameters which have been
found could be perceived or not. Hierarchical structuring of the musical parameters have
to be taken into account, although it is still unknown how they are structured.
Analyzing system
An analyzing system had to be found which is able to separate all different musical
parameters for computermusic. These are:
1. sections and subsections
2. streams of sound / layers (sounds which are separable from each other, develop
   differently etc.)
3. single sound parameters
4. direction of sound in space (right, left, moving etc.)
5. tension / movement in time
6. dynamic
7. length (of sections and sounds)
As a preliminary version the list can be extended with every new piece analyzed. If
parameters do not seem to be important or are unchanged throughout the piece (e.g.
dynamic), they simply can be left out.
Important are the single sound parameters (2). A sound will be described by some of the
attributes from the list of parameters. There are subdivisions like:
2.1. Spectral range bottom, bottom - middle, middle, middle - top, top, centered.
2.2. Relation of tone to overtones: harmonic, nonharmonic, sparse (few overtones, low
   amplitude), rich (many overtones, high amplitude)
2.3. Envelope: 2.3.1 beginning: attack, no attack (fade in, overtones building up, noises
   at start
   2.3.2 continuation: steady, beating
   2.3.3 ending: abrupt, decay (changing in overtones, lower amplitude),
   noises at end, transformation into new sound
2.4. Space Impression: reverberation, impression of distance etc. if it belongs to the
sound.
If further description is needed (e.g. if there are formants), it should be added. All
musical parameters except the single sound parameters are represented by graphical signs.
Handling the Analyzing System
As has been said earlier, analyzing will be done by reading the sonogram alongside
listening. While doing so several times, all parameter signs can be written into the
sonogram. It turned out to be most practicable to write all down later on an extra sheet.
All parameters should be listed for every section as a kind of abstract notation. Here,
the comparison of section lengths, sound qualities etc. will take place as well as the
interpretation of the results according to musicological questions.
Results
About 8 pieces of computer music have been analyzed with this system until now. Therefore,
some results can be given, although more analyzing must be done to be able to speak about
computer music in general. Analysis with sonograms helps to clear problems which cannot be
solved just by hearing analysis. That is especially the case, when sections overlap or
when similarity of sounds can be seen and better described literally. The lack of place
TCMC 552
allows just two short examples. The first is taken from Jonathan Harvey's "Ritual Melodies". The beginning of the piece is shown. The two sections are similar to each other in some points, e.g. length (11 sec.) and the kind of transformation of sound in the long layer. As analysis shows later, the first part of the piece is made up of eleven sections. With one exception, their duration is 11 sec. or it's multiple. Sometimes the sections overlap, than the duration includes the overlapping. The longest section of 55 sec. consists partly of the beginning sections in original or shortcuts (8sec.). The sections' overall duration will surely be not entirely perceived due to the capacity of short time memory. Most of the sounds seem to be in some relation to the ones of the beginning section. All in all, the construction principle of this piece's first part is quite evident. As time movement and dynamic built longer sections, this part (and the entire piece) seems to be carefully woven and held together, as students said after listening to the piece for the first time. Hence, some of the points may be perceived unconsciously.

The second example is taken from Trevor Wishart's "Vox 5". This piece is narrative, the sections and parts follow each other and there are only few similarities of sounds. It seems to follow the nature of a sounds, which in turn play with perceiving time, over the largest part and shows less construction in a mathematical sense. The part shown here is a constructed one, it is made up of three hierarchically interchanged repetitive patterns. For me it is interesting that this part is situated in the most difficult part of a piece: Around the end of the second third (in duration).

References

ICMC 553