Computational Music Archiving as Physical Culture Theory

Prof. Dr. Rolf Bader

Institute of Systematic Musicology
University of Hamburg
Neue Rabenstr. 13, 20354 Hamburg, Germany
R_Bader@t-online.de

ABSTRACT: The framework of the Computational Music and Sound Archive (COMSAR) is discussed. The aim is to analyze and sort musical pieces of music from all over the world with computational tools. Its analysis is based on Music Information Retrieval (MIR) tools, the sorting algorithms used are Hidden-Markov models and self-organazing Kohonen maps (SOM). Different kinds of systematizations like taxonomies, self-organazing systems as well as bottom-up methods with physiological motivation are discussed, next to the basic signal-processing algorithms. Further implementations include musical instrument geometries with their radiation characteristics as measured by microphone arrays, as well as the vibrational reconstruction of these instruments using physical modeling. Practically the aim is a search engine for music which is based on musical parameters like pitch, rhythm, tonality, form or timbre using methods close to neuronal and physiological mechanisms. Still the concept also suggests a culture theory based on physical mechanisms and parameters, and therefore omits speculation and theoretical overload.

2 BACKGROUND

2.1 KINDS OF MUSIC SYSTEMS

The aim of Systematic Musicology, right from the start around 1900, is the seek for universals in music, for rules, relations, systems or interactions holding for all musical styles of all ethnic groups and cultures around the world[61]. Music recordings and Phonogram Archives played a crucial role for establishing the field, as only after the invention of the Edison phonograph for recording music on wax cylinders[60] it was possible to compare music recorded by ethnomusicologists. The first of such archives was the Berlin Phonogram Archive established by Erich von Hornbostel and Carl Stumpf with a historical recording of a Thai phi pha orchestra at a visit in the Berlin Tiergarten 1900. Many recordings followed, like those of Jaap Kunst recording music of Indonesia[?].

2.1.1 TAXONOMY

One way of giving the endless variety of musics a system was the Hornbostel/Sachs classification of musical instruments[27] in align with the Sachs dictionary of musical instruments[54] as a detailed basis for such a classification. This classification system still holds today and is only enlarged by electronic musical instruments, whos development started in the second half of the 19th and was very prominent in the second half of the 20th century[51]. The success of this classification is caused by its classification idea of sorting musical instruments according to acoustical properties, namely their driving mechanisms. So instruments can be plucked, bowed, blown or struck, which produces similar timbres. This classification system is therefore showing a unity in the variety by comparing instruments. Therefore the field of Systematic Musicology in these days was called Comparative Musicology[59]. Taxonomies have been applied in many contexts in the field. A prominent example is that of musical styles forming feature lists[38] as applied to singing styles around the world[39].

2.1.2 SELF-ORGANIZATION

Classification is a hierarchical structure with global nodes followed by subnodes to differentiate plucked or blown instruments into many subcategories. This is only one kind of a system, and several others have been proposed. One such way is describing music as a self-organized system. The sounds produced by musical instruments are showing a very simple behaviour, the harmonicity of their overtone structures, only due to a very complex system of linear and nonlinear substructures, like turbulence in wind instruments or the bow-string interaction of the violin family[5]. The perception of music in the brain in neural networks is also a self-organizing process[35 37 67 68], where many neurons interact in nonlinear ways to result in simple outputs, the perception of timbre, rhythm, melodies or musical form. Therefore self-organization is a second system proposed to understand musics from all around the world in a more differentiated way[5].

2.1.3 PSYCHOACOUSTIC BOTTOM-UP PROCESS

Yet a third kind of system to understand music is often used in Musical Signal Processing and Music Information Retrieval (MIR)[29 16 17]. Here algorithms investigate the digital waveform of recorded music to retrieve information from this waveform, such as pitch, timbre, rhythm or other psychoacoustic parameters like roughness, brightness, density or the like. Classification of musical instruments is performed, as well as many other tasks like score following or Networked Music Performance[2]. Here the computer understands the music and can tell a player where he is in the score. Musicians around the world play together over the internet, and the task of the computer is to synchronize their playing, working around the restrictions of the speed of light, delaying the transmissions. This can be achieved by estimating the played music of a musician before it is actually played. Piano-roll extraction is yet another such task, where the computer understands typically piano music and prints a score from the sound wave file[10].

2.2 BASIC TYPES OF ALGORITHMS

Many of those tasks are realized by algorithms estimating a bottom-up approach to music retrieval, where the sound file is analyzed in terms of its spectrum or cepstrum at first which is then further processed using more complex algorithms to end up in the retrieval result[29]. To which extend those algorithms represent perception of music by humans, the processes physiologically present from the cochlear and the neural nuclei following up to the primary auditory cortex and beyond is not too much discussed. Indeed neural processing is self-orgnaization and therefore many of the algorithms are often only roughly related to perception[50 48 49]. Still this is not the main aim of such algorithms which can be seen as engineering solutions to a given task and often perform very well.

Other more complex algorithms have been proposed coming closer to human perception, like self-organizing maps of the Kohonen map type[35 37 67 68] . Here features of musical sounds are sorted in a map according to their similarities. Still as the map is organizing itself, no initial estimations are needed to decide about the way similarity may be measured, the features of the different sounds decide about this on their own. Also algorithms estimating the fractal dimension of sounds show considerable relatedness both to the processing of sound in humans as well as the perceptional sensation of musical density as a simple result to a complex computation[5].

Another kind of algorithms used for such tasks are Hidden Markov Models (HMM)[69]. Here the temporal development of events are predicted as the result of some hidden process. This process consists of the transition between a small amount of states, like musical pitches or musical instruments. The development of their appearance is modeled here as the probability of the transition of one state switching into another state, so e.g. one pitch followed by another one. As such transition probabilities are of statistical nature likelihoods describe the process and therefore the output is not fixed beforehand leaving space for arbitrariness. Still to which extend these models fit human perception is under debate.

Yet a totally different kind of algorithm describes music production of musical instruments. Physical Modeling is a set of methods to produce the sound of instruments by knowledge about the instrument geometries and the physical laws governing their vibration[9]. Several stages of complexity and simplicity exist here, from lumped models, digital waveguides or delay lines[53] to whole body geometries solving the differential equations governing the vibrations of plates, membranes or turbulent air flow[5 18]. These algorithms use the detailed geometry and solve the problems in a time-dependent manner, resulting in very realistic sounds and estimations of vibrating frequencies, transients or radiation. Using extreme parallel computation on an Field-Programmable Gate Array (FPGA) these geometries can be simulated in real-time[42 43]. Understanding the behaviour of instruments is often achieved in a bottom-up way by trying different models, adding or leaving out geometrical details, and from the comparison between the computed and measured sound decide about how the instrument works in detail. Other more simple models are able to show relations between musical instruments and instrument families more easily by starting with global estimations and adding necessary features in a topdown way[5].

Measurements of musical instruments are also a crucial part of the understanding the instruments and their relation one to another. Whole body measurement techniques, like microphone arrays[4], laser interferometry[41] or modal analysis? all give a detailed picture about the spatial and temporal development of the vibrations and transients of musical instruments. High-speed cameras and subpixel tracking analysis show the movement of strings or reeds[45]. This understanding leads to estimations of the global behaviour of instrument radiation, the role of different instrument parts in sound production, the use of materials like woods, hybrid- or metamaterials, the interaction of musical instruments with room acoustics or between instruments and players. Modern high-resolution methods of computer tomography (CT) give very detailed geometries of the instruments which can be used as input to physical modeling, showing details within the structures not accessible from their surfaces, or giving estimations of material parameters like density or Young's modulus. Therefore these methods are able to compare instruments and instrument families and give insight into building strategies and methods.

The different approaches to understanding music should at best all be used in a Computational Music and Sound Archive (COMSAR) as proposed here. Traditional phonogram archives only consist of recordings and their metadata, like the country they have been recorded, the musicians playing, the instruments in use etc.[13 65 66 46 12]. Still to address the aim of Systematic Musicology of finding universals in music[33], understanding its system, its production and perception need to use the analysis and analysis-by-synthesis tools discussed above. Including all these tasks is a tremendous effort, still all these fields nowadays show a high degree of specialization and are able to give detailed and robust results. Therefore to combine them together to an automatic analysis and search engine is straightforward.

2.3 COMSAR AS BIG DATA SOLUTION

Such automatic systems are needed in many fields. The endless amount of accessible music recordings via the internet, on CDs and in archives makes it practically impossible for a single researcher and even for research teams to perform these analysis by hand. Such a Big Data problem needs automatic tools for researchers to cope with. Additionally, the amount of methods and their complexity are so large that it is not feasible to have one researcher perform all tasks.

Also in terms of the 'buzz' the internet produces in terms of the endless variety of musics, research and consumer demands, search engines are needed to point researchers, musicians, listeners and music lovers to music they would hardly find otherwise. Such search engines need to be based on real musical parameters. Existing search engines for music used for mood regulation or work-outs sort music mainly in terms of its tempo and on its vitalizing properties. Still the musical styles of the world are so many and so differentiated that such simple parameters are not able to represent traditional music, all-in-a-box productions of musicians of ethnic groups in remote areas, boy groups in jungle regions combining their tradition with Western harmony or electronic music, free-improvisation, global Hip-Hop or Electronic Dance Music. Here much more di_erentiated algorithms are needed, and all of those mentioned above should be combined.

2.4 COMSAR AS PHYSICAL CULTURE THEORY

The COMSAR standard is not only to update traditional phonogram archives with modern methods and algorithms and not only about coping with Big Data. It is also the attempt to realize a culture theory based on physical reasoning.

As has been shown, both, musical instruments as well as neural brain networks are self-organizing systems[5]. They both are highly nonlinear and intensely coupled only to output a very simple behaviour. In terms of musical instruments this output is the harmonic overtone spectrum which would not be perfect without the self-organizing process at all. In terms of music perception and production, pitch, rhythm, melodies, musical form and other features are the results of self-organization and synchronization in the human brain.

Historically musical instruments have been developed and built by humans over at least the last forty thousand years according to the physiological and physical mechanisms the human brain and body is built of. The human voice, as well as animal vocalization are also self-organizing processes. They produce sounds which are not often found in non-living systems, namely the harmonicity of the partials. As voice is meant for communication, harmonic sounds are evolutionarily related to semantics. Therefore the semantics found in music need to be there because it is built-in the human auditory system reacting to harmonicity of the sounds.

Self-organization is the base of life compared to dead matter, it is turning non-living things to life. Its main issues are maintainance of life in a destructive world, difierentiation in parts to fulfill dificult tasks and the ability to assimilate in a changing environment. So building musical instruments as self-organizing systems means to make them similar to living systems. They exhibit behaviour we know only from animals or humans, harmonic tone production. Musical instrument builders have obviously decided to make this physical feature the core of musical instruments, and therefore the core of music as a cultural phenomenon. So the core of musical instruments is their self-organizing nature. We have built a music culture as artificial life by building musical instruments and perform on them.

A Physical Culture Theory is taking culture as a physical and physiological selforganization process building artificial life and therefore extending our life by inventing physical tools and processes which again work as self-organizing systems. Therefore the culture we build appears to us as a living system. The music speaks to us, the development of musical styles follow living behaviour, styles are born, live, die and are remembered, become legend. Musicians fuse with their instruments and experience them as having their own live, relate to them very similar to the way they relate to humans.

COMSAR, as implementation of many of the mechanisms and systems is therefore approaching music as a living culture, a selforganizing process. Of course it is only an approach yet, still extending the system in the future towards more and more precise algorithms and tools is only a matter of time.

Due to the difference of the algorithms discussed above, in the follow we give a deeper insight into main properties and research done in these fields, without being able to mention all of the works done here. MIR is already implemented in Phonogram Archives, physical modeling, microphone array techniques are not, and therefore work in ethnomusicology using these tools are given. Self-organizing maps as front end for search engines are discussed.

3 TOOLS AND APPLICATIONS FOR COMSAR

In 1978 Halmos, Kszegi, and Mandler coined the term Computational Ethnomusicology for using MIR also in regard to non-Western music[23]. Since then many tools and applications have been suggested to retrieve, sort and understand musical content from sound.

3.1 BOTTOM-UP MIR TOOLS

In early attempts to extract musical parameters from sound, simple multi-line textures consisting of two voices were considered, which must not have overlapping over-tones[44]. More modern approaches include percussion sounds, meter and rhythm estimation[30] and are designed for the analysis of harmonic as well as percussive instruments[24], including psychoacoustic knowledge[31], or using different approaches in the matrix domain[14]. Similarity matrices of spectral features, like Mel Frequency Cepstral Coefficients (MFCC), amongst others, have also been proposed to relate parts of a piece, like verse or chorus[20]. Singular Value Decomposition (SVD) has also been employed in this context[21]. Recently TARSOS, a platform to extract pitch information from sound using 1200 cent per octave has been developed[62] to suit demands in Computational Ethnomusicology[23].

The first task to accomplish in analysing audio signal will invariably be the detection of the onset of any given signal event (see[15] for a review). The approaches employed here range from measurement of strong amplitude raise and phase differences, to fluctuation estimation. It appears that the choice of the onset detection algorithm depends on the type of sound to be analyzed. For percussive sounds, measurement of amplitude raise is sufficient, yet for fusing tones, like piano or violin sounds, measurement of fluctuations seems more promising. As fluctuations on a phase level seem to include both to a great extend, the approach favoured by the applicant and his staff is therefore based on a Modified Modulation Algorithm[2].

A second task is the estimation of pitch, often referred to as f0-estimation, i.e. detecting the fundamental partials of a given harmonic spectrum[25 29]. An approach used in the context stated in this proposal is based on algorithms like Autocorrelation Functions, but furthermore employs Correlogram Representation for f0 estimation in multi-line textures[11,6]. This robust method allows estimation of harmonic overtone structures within very short time frames. Additionally, to estimate if a piece is single- or multi-line, the Fractal Correlation Dimension is appropriate, as the integer dimension number thus obtained constitutes the amount of harmonic overtone series present in a given musical sound[22 5].

3.2 SELF-ORGANIZING TOOLS

Representation of the results of an analysis for use in IR algorithms has been proposed in several ways. COMSAR uses self-organizing maps (SOM), Hidden-Markov Models (HMM), and correlation matrices, all based on the extracted data.
Self-organizing Kohonen maps have been proposed for pitch and chord mapping[37,36], and for sound level assessment[35], for a review see [5]. This method has also been successfully applied to related fields, such as speech estimation[47], and soundscape mapping. Here the feature vector extracted by the MIR algorithms, consisting of pitch contour, spectral centroid, uctuations, inharmonicity, etc., is fed into an Artificial Neural Network within a defined training space. After the training process, any such system should be able to identify new feature vectors by itself and will therefore be able to define a parameter space for these features for all of the analysed archival assets, and will be able to detect structural similarities on a best-estimation basis.

3.3 HIDDEN-MARKOV MODEL

A complementary approach to be employed is the implementation of the Hidden-Markov-Model (HMM), used for stochastically estimating transition probabilities between hidden states, which, performed consecutively, results in an event series, as present with both, musical rhythm and melody. These models have been used extensively for musical applications[1 2]. The Markov model consists of musically meaningful states. So when representing, for example, a multi-line rhythm, these states could be bass drum, snare drum, hi-hat, tom-tom, etc. These are mathematically represented as a Mixed Poisson distribution.

Additionally, a transition matrix between these hidden states will be calculated using an Estimation-Maximization (EM) algorithm[69]. Both, the Poisson distribution and the transition matrix determine the musical parameters, rhythm, melody and texture. This representation may then be compared to all previously analysed assets in the ESRA database, again forming a state-space, detecting similarities, relate objects, etc.

4 COMSAR ARCHITECTURE

A MIR-based data infrastructure and classification scheme is to be implemented within the framework of the ESRA database currently under development to be able to categorize the database content in regard to basic musical parameters derived from the digital audio data stream.

The three main musical parameters which are treated using the MIR analysis described in this proposal are pitch (melody, texture), rhythm (single- and multi-line), and timbre (single- and multi-line). The MIR structure has two main threads, the timbre thread (TT) and the pitch thread (PT). As TT deals always with the whole sound information, PT performs a pitch extraction from the sounds and proceeds with pitch information only.

4.1 TIMBRE THREAD (TT)

The first step in TT is a segmentation of the audio file in terms of onset detection (OD). Here, two main methods are used, the fluctuation method for fusing tones[2] and a simple amplitude model for percussive onsets[29]. From the segments three MIR estimations are performed: a Timbre Thread Rhythm (TTR), a Timbre Thread Timbre Multi-line (TTM) and a Timbre Thread Timbre Singleline (TTS).

4.1.1 TIMBRE THREAD RHYTHM (TTS)

TTR does take the sound played by several instruments as one; it does not attempt any splitting of compound sounds into individual instrument sounds. As discussed above, retrieving individual sounds of musical instruments from a multi-instrumental recording is theoretically impossible, because of the fact that no clear association with all partials of harmonic pitch structures can be assumed from the sound alone without any further knowledge. Still to be able to deal with more complex rhythms, in the PT section (see below) a multi-line estimation is performed to detect the most probable events without the need to extract the sounds perfectly.

Within the TTS, for each segment a spectral centroid is calculated as the most prominent parameter of timbre perception. The list of centroid values of the onsets found is then fed into a Hidden-Markov Model (HMM), using a Poisson Mixture Model (PMM). The results of the HMM are the parameters of the PMM, which represent the rhythmical structure of the centroid values of the onsets. This PMM, as well as the Transition Probabilities (TP) are calculated for all objects in the database and a correlation matrix between all PMMs and TPs is calculated to relate the di_erent rhythm PMM structures in terms of similarity.

4.1.2 TIMBRE THREAD TIMBRE MULTILINE (TTM)

As discussed in the Pitch Thread (PT) section below, it is estimated if a given recording contains multi-line or single-line melody (this may also be judged aurally and used as additional, external input). Additionally, the fractal correlation dimension D of a given piece is calculated for adjacent sound sections of 50ms. If 100ms after the initial transient D >= 2.0, the sound has more than one harmonic overtone structure and therefore is considered multi-line. Within this definition, all percussion objects are multi-line, too. This is reasonable also if only one drum is played. If the piece is found to be multiline in nature, the TPM algorithm estimates a feature vector of each segment provided by the onset detector, using spectral centroid, uctuations within the steady-state of the sound, amount of chaoticity of the initial transient, and other related features found with timbre perception of multi-dimensional scaling events. These features are calculated for adjacent times within each segment to end in a multidimensional trajectory of the sound development, as found crucial to explain nearly endless possible sounds within a low-dimensional timbre space, by adding the temporal development of the sound within this space.

This feature vector is then used as input to a self-organizing Kohonen map. After training, this map constitutes a twodimensional representation of the objects in the database. All segments of all objects are then fed into the map, where the neuron with maximum similarity between the given segment and this neuron positions the segment within the map. Therefore, segments or objects can be estimated for similarity from the trained map.

4.1.3 TIMBRE THREAD TIMBRE SINGLELINE (TTS)

If a piece is found to be single-line all through, as discussed above, the same procedure is performed, training a Kohonen map with the feature vector of the given sound. Again the trained map is then able to relate all segments and all pieces, and give similarity judgements. The reason why the single- or multi-line cases are separated is to have one map which is able to classify single instruments alone, while the other is able to deal with orchestrated multiinstrument sounds. So if a musical instrument is to be judged in terms of similarity, the TTS can be used. Another reason is the problem of dealing with the different pitches of the sounds. The TTS map will classify both, pitch and timbre. As pitch is the most prominent factor in musical instrument similarity judgements, the map will have different regions for different pitches. Then within each region the differentiation in terms of timbre is present. This is automatically performed by the map. Still it is necessary, as one instrument may sound considerably dfferent within different registers. Differences in articulation within one pitch region will again be met by the differentiation of the map within the pitch region of the sound investigated. This cannot be done with multi-line sounds, as here virtually endless possibilities of pitch combinations can be present.

4.2 PITCH THREAD (PT)

PT is representing a piece on the score level, although of course it also needs to start from the recorded sound. So, first PT performs a pitch extraction, both single- and multi-line. Two main algorithms are used here, the correlogram for multi-line and the autocorrelation function for single-line sounds. The correlogram is detecting whole overtone structures, which are related to pitch, and finds the basic frequency for it within small time frames of about 20ms. As it also displays multiple harmonic series, the pitches of different instruments can be detected with high frequency resolutions. If a piece is singleline, this algorithm can be used, too. Additionally, an autocorrelation estimation of small time frames of again about 20ms is performed, adding information to the correlogram. The result is a temporally and spectrally high resolution function of the harmonic overtone series, the pitches, over time, for the whole object. From this pitch texture, again three musical parameters are calculated, the Pitch Thread Rhythm (PTR), the Pitch Thread Texture (PTT), and the Pitch Thread Melisma (PTM).

4.2.1 PITCH THREAD RHYTHM (PTR)

For PTR, from the pitch texture, note onsets are calculated to end up in a musical score. This score consists of pitch events, both in terms of Western pitch classes as well as in term of their microtonal precisions up to 1200 cent per octave over time. As with the TTR, a Hidden-Markov Model (HMM) is used with a Poisson Mixture Model (PMM) which has as many hidden states as are di_erent pitches appearing in the given object. The PMM and the Transition Probability (TP) calculated by the HMM then represents the objects, which can therefore be related in terms of similarity.

4.2.2 PITCH THREAD TEXTURE (PTT)

Here, from the pitch texture again the onsets are calculated, again to end in a score. For PTT, this score itself is used to correlate the objects in terms of their similarities. PTT is meant for objects with no glissando, where the pitch texture holds the main information of the object.

4.2.3 PITCH THREAD MELISMA (PTM)

Contrary, PTM is meant for objects in which pitch changes are very important, in terms of vibrato, glissandi, etc. Again after performing onset detection, the objects are divided into small segments for each played note. Still, all detailed pitch information over time within the segments is preserved here. Again, a Kohonen map is used to represent possible ornamentations, melismata, glissandi, or vibrato. So here special ornamentations or melismata can be compared, rather than whole objects themselves.

4.3 SUMMARY OF THREADS

After performing all these analysis of all objects, six parameters result, which estimate the basic musical parameters timbre, pitch, and rhythm for each piece:

Timbre

Kohonen map of multi-line timbre (TTP)
Kohonen map of single-line timbre, musical instrument sounds (TTH)

Pitch

Score for multi- and single-line objects (PTT)
Kohonen map of melisma for all segments of all objects (PTM)

Rhythm

Hidden-Markov Model of multi-line fused sounds based on sound level (TTR) Hidden-Markov Model of multi- and single-line objects based on pitch level (PTR) With these six models all objects within the archive can be compared in terms of all the sub-features present in the models. Also new objects or sounds can be compared to all the existing features in the model. The similarities proposed by the algorithms then need to be judged by listeners and experts.

5 INCLUDING MUSICAL INSTRUMENT MEASUREMENTS AND MODELING IN COMSAR

Organology was part of understanding and systematizing music as part of the Hornbostel/ Sachs classification. Musical instrument dictionaries like that of Curt Sachs (see above) mention and describe thousands of musical instruments from all over the world. Features like their origin and use in the musical culture, the material they are built of or the building process are documented. These dictionaries are a useful source when it comes to identifying instruments collected in the field and for giving information about their content.

Still research is way ahead in terms of the acoustics, properties and building processes of musical instruments. The basic principles of how musical instruments vibrate and radiate sound have extensively been studied (for reviews see[3,19,?,?]). Many musical instruments have been investigated in great detail, mainly those of the West, but also many others all over the world. The materials used are known in terms of their material parameters like Young's modulus, density or internal damping. The building process of many instruments have been described not only as plain craftsmanship but also in terms of the acoustical and musical function these processes have been motivated by.

Several theoretical frameworks on the acoustical properties of musical instruments have been developed over the last decades making it possible to classify them not only in terms of their driving mechanisms, like Hornbostel/ Sachs have done, but by their physical mechanisms and features. The Impulse Pattern Formulation (IPF) considers musical instruments as working with short impulses caused by one part of the instrument, e.g. the force of a string acting on a body or the pressure impulse of wind instruments produced at the players mouth. These impulses are transferred to other parts of the instruments, like top- and back plates, rims and ribs, are filtered and return to its origin. In a self-organizing process this system starts with a transient phase which is complex and chaotic only to organize itself after a short time of maximum 50 ms to end in a harmonic overtone sound radiation[5]. Other proposals are that of a nonlinear driving generator and a linear resonator which interact, producing sound, or that of phase-locking of different partials (for a review see[5]).

Also the geometry of the instruments are known in great detail. High-resolution Computer Tomography (CT) scans of whole instrument geometries display the instruments with resolutions of a fraction of a millimeter. From these results material properties can be derived, like density, speed of sound, diffraction or internal damping. Therefore the rough estimations of geometrical data have been replaced by detailed and precise measurements.

The radiation of musical instruments have also been measured extensively, as only during the last years technological advances have made it possible to recored single instruments with microphone arrays consisting of up to 128 microphones when recording single sounds and up to several thousand microphone positions when recording multiple sound instances (for a review see[4]). Some techniques allow the back-propagation of the radiated sound to the radiating surface, the musical instrument geometry. This means a measurement of the instrument vibration all over its geometry within and therefore a measurement of the internal vibrational energy distribution, the role of geometrical parts to the acoustical output, or an estimation the radiation of the instrument at any place in a performance space.

Physical modeling of musical instruments have also been performed extensively over the last decade (for a review see[7,9,3]). Here the differential equations governing the vibration of the instruments are used with a geometrical model of the instrument to make this virtual instrument vibrate in silico. High time and spatial resolution allows the precise modeling of the instrument and the production of a sound very close to the original sound of the instrument. By changing the mathematical model the role of different kinds of vibrations, couplings or instrument parts can be shown. By changing the geometry or the material parameters the instruments can be understood in terms of why which geometry is used and how changed here would change the sound of the instrument. The use of geometrical changes or alternative materials can be tested before building the instrument, a property needed nowadays as climate change forces new wood species to be planted e.g. in Europe.

All these contemporary features of musical instrument research, models and experimental setups need to be part of a phonogram archive standard in the near future. Two examples are given below, one about the use of microphone arrays to measure the acoustical behaviour of a lounuet, a New Ireland friction instrument, and the use of paste to tune the membrane of a Burmese pat wain. Both examples show the clarification of ethnomusicological questions about the instruments. A modern way of using microphone-array measurements, high-resolution CT scans or physical modeling need to be developed in the future and faces several issues, like server space, computational capability of servers or the retrieval of the information necessary to built such models. Still solutions are around for all aspects so that the implementation of these features to COMSAR is mainly a matter of time and energy and not a fundamental problem.

The advantages of such methods are many. Instrument builders could go online and look for the instrument they build, or for similar ones, understand more about its vibration, radiation or the inuence of several parts or materials on the sound, include changes of the material or parameters online and listen to the resulting sound of such a new instrument instantaneously. They could then decide to built such an instrument or thing of different changes. They could also be inspired by similar instruments, their building process, materials or sounds and decide to try new instruments similar to their traditional one.

Researchers could estimate how important different aspects of the instrument for sound production are. Some building processes of instruments are sound decisive, some may be needed in terms of rituals used for the building process, and some might be pure myths, traded by tradition rather than by a sound idle and unnecessary or even unwanted today. In the history of instrument building of Western instruments, like violins, pianos, trumpets or guitars, many of these myths have been identified over the years and from an ethnomusicological standpoint it is important to know which stories are true and which are not.

Also from an educational point of view such a system would be highly attractive to young people interested in the music of the world and used to use the internet, search engines and simple music production systems to be creative. Such tools could be rated as 'cool' and contemporary and therefore be used with the by-effect of making them understand the principles of musical instruments and their use.

In terms of replacing wood and other natural material becoming scarce nowadays with artificial materials, metamaterials or the like, such tools would be highly welcome, too, as everybody could try online how such changes would effect the sound and if they might be used.

Many other applications are to be expected due to an inclusion of these techniques in phonogram archive standards in the near future.

6 CONCLUSION

A Computational Music and Sound Archive does not only fit the needs of sorting and analyzing music automatically in times of Big Data and digital acscessibility of music. It also is a way to understand music and culture in terms relating cultures and ethnic groups rather than stress the dissimilarities between them. It is also objective in a way to omit cultural bias and view. Therefore it suggests a view on music as a complex system rather than sorting it in genres or styles. Such a Physical Culture Theory is therefore both able to cope with the complexity of todays reality as well as suggesting a new and fresh look on music in the world.

7. REFERENCES

[1] Aucouturier, J.-J. and M. Sandler 2001: Segmentation of Musical Signals Using Hidden Markov Models, in Proc. of the Audio Engineering Society 110th Convention.

[2] Alexandraki, Ch. and R. Bader 2013: Real-time Concatanative Synthesis for Networked Musical Interactions. J. Acoust. Soc. Am. 133, 3367, 2013.

[3] Bader, R.: Springer Handbook of Systematic Musicology. Springer, Berlin, Heidelberg 2018.

[4] Bader, R.: Microphone Array. In: T. Rossing (ed.): Springer Handbook of Acoustics. 1179-1207, 2014.

[5] Bader, R. 2013: Nonlinearities and Synchronization in Musical Acoustics and Music Psychology. Springer Series Current Research in Systematic Musicology, Vol. 2. Heidelberg. Springer.

[6] Bader, R. 2011: Buddhism, Animism, and Entertainment in Cambodian Melismatic Chanting smot. In A. Schneider and A. von Ruschkowski (eds.): Hamburg Yearbook of Musicology 28.

[7] Bader, R. & Hansen, U.: Acoustical Analysis and Modeling of Musical Instruments Using Modern Signal Processing Methods. In: Handbook of Signal Processing in Acoustics. D. Havelock, M. Vorlnder, S. Kuwano (ed.). Springer, 219-247, 2008.

[8] Bader. R.: "Finite-element calculation of a bass drum," J. Acoust. Soc. Am. 119, 3290, 2006.

[9] Bader, R.: Computational Mechanics of the Classical Guitar. Springer Oktober 2005.

[10] Benetos, Emmanouil & Dixon, Simon: Multiple-instrument polyphonic music transcription using a temporally constrained shift-invariant model. J. Acoust. Soc. Am. 133 (3), 1727-1741, 2013. 15

[11] Chai W. and B. Vercoe 2003, Structural Analysis of Musical Signals for Indexing and Thumbnailing, in Proc. ACM/IEEE Joint Conference on Digital Libraries.

[12] Copeland, P. 2008. Manual of Analogue Sound Restoration Techniques. London. The British Library.

[13] Elschek, Oskr et al. 2001. Digitizing world music. Digitalisierung von Weltmusik. Special issue, Systematische Musikwissenschaft/Systematic Musicology Vol. VII, no. 3.

[14] Dessein, A., A. Cont and G. Lemaitre. 2010. Real-time polyphonic music transcription with non-negative matrix factorization and beta-divergence, In Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR), pp. 489494.

[15] Dixon, S. 2006. Onset detection revisited. In Proc of the Int Conf on Digital Audio E_ects DAFx06, pp. 133137.

[16] Downie, J. Stephen 2003. Music information retrieval. Annual Review of information science and technology 37, 295-340. (Available from http://musicir.org/downie mir arist37.pdf).

[17] Fingerhut, M. 2004. Music information retrieval, or how to search for (and maybe _nd) music and do away with incipits. Proc. IAML IASA Joint congress, Music and Multimedia, Oslo, August 2004. (Avalaible from http://www.ismir.net/admin/ismirbookle.pdf)

[18] Fischer, Jost: Numerical simulations of the turbulent flow and the sound field of the Turkish ney end-blown ute.J. Acoust. Soc. Am. 141, 3960, 2017.

[19] Fletcher, Neville & Rossing, Th. D.: Physics of Musical Instruments. Springer, Heidelberg 2000.

[20] Foote, J.: Automatic Audio Segmentation Using a Measure of Audio Novelty, in Proc. ICME'00, 2000.

[21] Foote, J., M. Cooper 2003 Media Segmentation Using Self-Similarity Decomposition, in Proc. Of SPIE Storage and Retrieval for Multimedia Databases, Vol. 5021.

[22] Gibiat, V. & Castellengo, M. 2000: Period Doubling Occurences in Wind Instruments Musical Performance. Acustica 86, 746-754.

[23] Gmez, E., Herrera, P. & Gmez-Martin, F. 2013: Compuational Ethnomusicology: perspectives and challenges. J. New Music Research 42:2, 111-112.

[24] Goto, M. 2001a. An Audio-based Realtime Beat Tracking System for Music With or Without Drum-sounds, Journal of New Music Research 30(2): 159-171.

[25] Goto, M. 2001b. A predominant-f0 estimation method for real-world musical

audio signals: MAP estimation for incorporating prior knowledge about f0s and tone models, In Proc Workshop on Consistent and reliable acoustic cues for sound, 1-4.

[26] Harwood, D. L. 1976. Universals in Music: a Perspective from Cognitive Psychology. Ethnomusicology, 20, 521-533.

[27] Hornbostel, Erich M. v. & Sachs, C.: Systematik der Musikinstrumente. Ein Versuch. Zeitschrift für Ethnologie 46, 4-5, 553- 590, 1914.

[28] International Association of Sound and Audiovisual Archives Technical Committee (eds.) 2005. IASA-TC 03 Standards, Recommended Practices and Strategies: the Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy. http://www.iasaweb.org/IASA TC03/TC03 English.pdf. as of 2013-09-13.

[29] Klapuri, A, & M. Davy: Signal Processing Methods for Music Transcription. Ed. Anssi Klapuri and Manuel Davy. Signal Processing. Springer-Verlag New York Inc, 2006.

[30] Klapuri A: Signal Processing Methods for the Automatic Transcription of Music. PhD Dissertation, Tampere University of Technology, 2004.

[31] Klapuri, A.: Sound onset detection by applying psychoacoustic knowledge. 1999 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, ICASSP99 Cat No99CH36258 6, 3089-3092, 1999.

[32] Kolinski, M. 1957. The Determinants of Tonal Construction in Tribal Music. The Musical Quarterly, 43(1), 50-56.

[33] Kolinski, M. 1978. The Structure of Music: Diversi_cation Versus Constraint. Ethnomusicology, 22(2), 229-244.

[34] Kolinski, M. 1982. Reiteration Quotients: a Cross-Cultural Comparison. Ethnomusicology, 26(1), 85-90.

[35] Kostek, B. 2005: Perception-based data processing in acoustics. Applications to Music Information Retrieval and Psychophysiology of Hearing. Springer.

[36] Leman, M 1995.: Music and schema theory. Berlin, Springer.

[37] Leman, M. and F. Carreras 1997: Schema and Gestalt: Testing the hypothesis of Psychoneural Isomorphism by Computer Simulation. In: Marc Leman (ed.): Music, Gestalt, and Computing. Studies in Cognitive and Systematic Musicology. Springer, Berlin, 144-168.

[38] Lomax, A., and N. Berkowitz 1972. The Evolutionary Taxonomy of Culture. Science 177(4045), 228-239.

[39] Lomax, A. 1976. Cantometrics: an Approach to the Anthropology of Music. Berkeley: University of California Extension Media Center.

[40] Moelants, D. et al. 2007. The Problems and opportunities of content-based analysis and description of ethnic music. International Journal of intangible heritage 2, 57-67.

[41] Moore, Thomas: Measurement Techniques. In: Bader R. (eds) Springer Handbook of Systematic Musicology. Springer, Berlin, Heidelberg, 81-103, 2018.

[42] Pfeie Florian: Real-Time Signal Processing on Field Programmable Gate Array Hardware. In: Bader R. (eds) Springer Handbook of Systematic Musicology. Springer, Berlin, Heidelberg, 385-417, 2018.

[43] Pfeie, Florian, Bader, Rolf: Real-Time Finite-Di_erence Method Physical Modeling of Musical Instruments Using Field-Programmable Gate Array Hardware, J. Audio Engineering Society 63 (12), 1001-1016, 2015.

[44] Pisczcalski M. and B. Geller 1977. Automatic Music Transcription, Computer Music Journal, 1(4): 24-31.

[45] Plath, Niko: High-speed camera displacement measurement (HCDM) technique of string vibrations,Proceedings of the Stockholm Music Acoustics Conference, 188-192, 2013.

[46] Proutskova, P. 2007. Musical Memory of the world Data infrastructure in ethnomusicological archives. Proc. of the 8th

Intern. Conf. On Music Information Retrieval (ISMIR), Vienna, Austria 2007.

[47] Rabiner L. R. and B. H. Juang 1993. Fundamentals of Speech Recognition, Prentice Hall Signal Processing Series

[48] Rohrmeier M., Pearce M.: Musical Syntax I: Theoretical Perspectives. In: Bader R. (eds) Springer Handbook of Systematic Musicology. Springer, Berlin, Heidelberg, 473-486, 2018.

[49] Pearce M., Rohrmeier M.: Musical Syntax II: Empirical Perspectives. In: Bader R. (eds) Springer Handbook of Systematic Musicology. Springer, Berlin, Heidelberg, 487-505, 2018.

[50] Rohrmeier, M. A., & Cross, I. (2014). Modelling unsupervised online-learning of artificial grammars: Linking implicit and statistical learning. Consciousness and Cognition: An International Journal, 27, 155-167, 2014.

[51] Ruschkowski, Andre: Elektronische Klänge und musikalische Entdeckungen [Electronic Sounds and musical discoveries], Reclam, Ditzingen 1998.

[52] Rzeszutek, T., Savage, P. E., & Brown, S. 2012. The Structure of Cross-Cultural Musical Diversity. Proceedings of the Royal Society B: Biological Sciences, 279, 1602-1612.

[53] Scarvone, Gary: Delay-Lines and Digital Waveguides. In: Bader R. (eds) Springer Handbook of Systematic Musicology. Springer, Berlin, Heidelberg, 259-272, 2018.

[54] Sachs, Carl: Real-Lexikon der Musikinstrumente : zugleich ein Polyglossar für das gesamte Instrumentengebiet. Bard, Berlin 1913.

[55] Sachs, C.: Geist und Werden der Musikinstrumente. [Spirit and Becoming of Musical Instruments.] Berlin: Reimer, 1929.

[56] Sadie, S.: The New Grove Dictionary of Musical Instruments. Vol. 2. Macmillan Press Limited, 1984.

[57] Savage, P.E., Merritt, E., Rzeszutek, T., & Brown, S. 2012. CantoCore: a New Cross-Cultural Song Classification Scheme. Analytic Approaches to World Music, 2(1), 87-137.

[58] Scavone, Gary: Delay-Lines and Digital Waveguides. In: Bader R. (eds) Springer Handbook of Systematic Musicology. Springer, Berlin, Heidelberg, 259-272, 2018.

[59] Schneider, A. 2001. Sound, Pitch, and Scale: from Tone Measurements to Sonological Analysis in Ethnomusicology. Ethnomusicology, 45(3). 489-519.

[60] Schneider, A. 2006 Comparative and Systematic Musicology in relation to Ethnomusicology: A Historical and Methodological Survey. Ethnomusicology, 50(2), 236-258.

[61] Schneider, A. 2018 "Systematic Musicology: a historical interdisciplinary perspective". In: R. Bader (ed.): Springer Handbook of Systematic Musicology, 1- 24, Springer Heidelberg, Berlin.

[62] Six, J., O. Cornelis and M. Leman 2013: Tarsos, a Modular Platform for Precise Pitch Analysis of Western and Non-Western Music. J. New Music Research 42:2, 113-129.

[63] Tenzer, M. (Ed.). 2006. Analytical studies in world music. New York: Oxford University Press.

[64] Tenzer, M., & Roeder, J. (Eds.). 2011. Analytical and cross-cultural studies in world music. New York: Oxford University Press.

[65] The National Recording Preservation Board of the Library of Congress (eds.) 2010. The State of Recorded Sound Preservation in the United States: a National Legacy at Risk in the Digital Age. Washington, D.C. Council on Library and Information Resources and The Library of Congress.

[66] The National Recording Preservation Board of the Library of Congress (eds.) 2012. The Library of Congress National Recording Preservation Plan. Washington, D.C. Council on Library and Information Resources and The Library of Congress.

[67] Toiviainen, P. (Ed.) 2009. Musical similarity [Special issue]. Musicae Scientiae, 13(1 suppl).

[68] Toiviainen, P., and Eerola, T. 2001 A method for comparative analysis of folk music based on musical feature extraction and neural networks. In H. Lappalainen (Ed.), Proceedings of the VII International Symposium of Systematic and Comparative Musicology and the III International Conference on Cognitive Musicology (pp. 41-45). Jyvskyl: University of Jyvskyl.

[69] Zucchini, W. and I.L. MacDonald 2009: Hidden-Markov Models for Time Series. An Introduction using R. Chapman & Hall.