

The Musical Communication Chain and its Modeling

Xavier Serra
Audiovisual Institute, Pompeu Fabra University
Barcelona, Spain
<http://www.iaa.upf.es>
email: xserra@iaa.upf.es

[published in Assayag, G. et al. Ed. 2002. *Mathematics and Music. A Diderot Mathematical Forum.* Springer Verlag.]

1. Introduction

We know that Music is a complex phenomenon impossible to approach from any single point of view; thus, most scientific attempts to understand the music process do not make justice to its amazing richness. However, this cannot prevent us from trying to explain and to formalize some aspects, thus contributing to its understanding and at the same time offering new tools to create and enjoy new music works, thus participating in enriching that same marvelous complexity.

In his book *Elements of Computer Music* (Moore, 1990), Richard Moore presents a view of the musical communication chain that despite being based on a traditional music conception, it is a very useful starting point for discussing many relevant issues connecting computers and music. Starting from that view, in this article we will mention some of the active research areas and present topics that are still very much open to be looked into. No attempt is made to present a comprehensive overview of the Computer Music field and its related disciplines. A part from the book by Moore there are a few books that can give the reader an overview of the field (Roads, 1996; Dodge and Jerse, 1996).

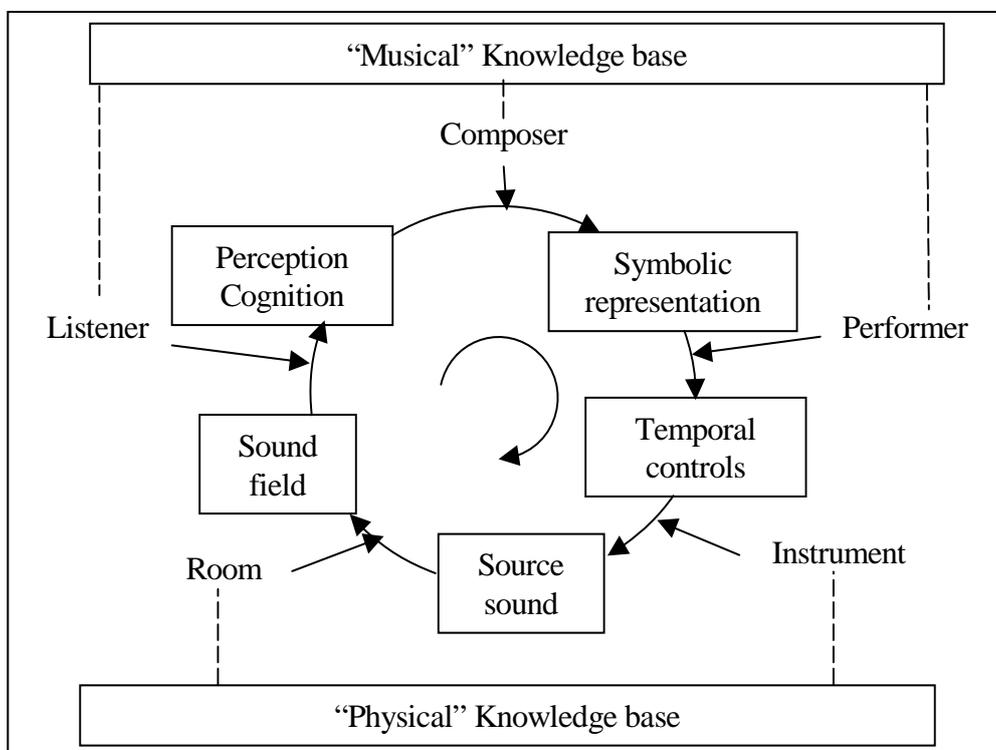


Figure 1: Diagram of the Musical Communication chain proposed by Moore.

2. The Communication Chain

The musical communication chain proposed by Moore, and shown in the figure, is a loop of interconnected signals (data) and processes (transformations of the signals) that encompass all the elements involved in the making, transmission and reception of music. As shown in the diagram, this loop is the result of the encounter of two knowledge bases, the musical one, basically mental and fundamented on our cultural tradition, and a physical base, in which the laws of physics have a much greater influence.

Starting at the top of the loop we can see that from the perceptual inputs and a given personal music background, the composer is able to create a symbolic representation that expresses a musical idea. From such symbolic representation and by sharing some of the same music background, a performer is then capable of producing gestures, or temporal controls, to drive a music instrument. Such an instrument is a crafted physical object that can produce an air vibration, sound source, from the performer's gestures. The sound source produced by the instrument is then propagated in a room and a sound field is created which a listener then perceives. Finally, the listener processes the acoustic signal that enters the ear by using previous perceptual experiences, thus having a perceptual and cognitive experience. The composer closes the loop by using his/her own perceptual and cognitive experiences in the musical creative decisions.

This description is a traditional view of the music communication chain and it is clear that the scientific and technological developments of the second half of the 20th century have had a great impact in this chain. An important alteration caused by these developments has been the flexibilization of some of the processes, since now they can not only be undertaken by human beings or mechanical devices, but also by electronic machines or even by software simulations. For example, a physical room, might be completely skipped in the communication chain, since we can listen to a performance on headphones, or a performer might also be leaved out since the composer can produce sounds directly with a computer without the traditional physical interaction performer-instrument. But the most important impact has been the incorporation of new creative possibilities for the composer, performer and instrument builder, many of which come from the potential of creating computer models for specific processes involved in the chain.

Despite all the possible shortcomings and simplifications of the chain, its discussion will give us useful insights that are very much applicable to the current developments in the area of Computer Music. In fact, we are particularly interested in discussing this chain in the context of the current scientific and technological developments, viewing how they affect, and maybe redefine, the communication chain. Let us now go into some detail through the different parts of the chain.

3. Composer

Traditional music composition is the process of producing symbolic representation of musical thought. Even though musical thought is difficult, or even impossible, to formalize, there is a long tradition of formalization of the compositional process and of its symbolic output. From this tradition it has been a quite natural step to use the computer as an assistant in the compositional process. Developing formal composition algorithms; modeling some specific compositional tasks and devising tools for helping the composer (Cope, 1991; Hiller and Isaacson, 1959; Loy, 1989; Schillinger, 1948; Xenakis, 1971).

The idea of automated, or algorithmic, composition is as old as music composition itself. It generally refers to the application of rigid, well-defined rules to the process of composing music. Since composers always follow some rigid rules and structures, we could say that classical compositions are also algorithmic compositions. However, we normally restrict the term of

algorithmic composition to the situations when there is minimal human intervention. The techniques used in the automated composition programs (Brown, 1998) cover a wide spectrum of approaches coming from such different domains as law (rules), mathematics (mathematical functions), psycho neurology (connectionism), and biology (generative processes).

- *Rule-based.* The most common of the early algorithmic systems were those which applied rules, of counterpoint for example, to musical choices. These systems rely on the “wisdom” of those who design the rules. Rule-based systems can be extremely complex but always rely on the specification of musical “heuristics” by the programmer. Rules can be derived from analysis of previous musical works or from other formal structures such as linguistic grammars, and mathematical formulae.
- *Mathematical Functions.* The direct use of mathematical functions has also been a fruitful source of inspiration for the development of algorithmic system. For example using probability functions, when choices between various options are weighted, or more recently with the use of equations derived from the Chaos theory or Fractal structures.
- *Connectionist.* A connectionist system, such as a neural network, is “trained” by being exposed to existing music, at which time it alters the characteristics of the connections between its nodes (neurons) so they reflect the patterns in the input. After training the connectionist system it can be seeded and then produce an output based on its learned patterns.
- *Generative Processes.* Generative systems are based on theories of genetic evolution. The basic idea is that a melody, or a more complex music structure, can “grow” similarly to the evolutionary development of a life form. The oldest form of evolutionist computer algorithms are “cellular automata”. Another common class of evolutionary algorithms are those based on “genetic algorithms”, which model the splitting, mutation, and recombination of genes. For example, within a gene pool of notes new children can be created by taking the pitch from one parent and the duration of another. Mutations are created by randomly changing the values of the data and a melody is generated by selecting notes from the pool of notes.

Every single composer has a particular way of thinking about music and thus it has been difficult to come up with software applications that can be used by more than a few composers or compositional styles. The most successful systems are either the ones that are very specific, thus, are useful for a very precise task, or the ones that do not assume any given aesthetic or mental process and give a lot of freedom to the composer to make his/her own music model. This last group of systems are, essentially, computer languages that support basic music and sound constructs from which the user can express his/her own musical thought.

4. Symbolic Representation

The traditional symbolic representation of classical western music, Common Practice Notation, CPN, is a highly encoded abstract representation of music that lies somewhere between instructions for performance and representation of the sound. It has been an excellent way to communicate between composers and performers for several centuries and it assumes that performers and composers share a common music tradition, thus the performers can infer many of the non-written performance instructions from that tradition. The usefulness of this representation breaks down as soon as the shared tradition does not exist, as in some contemporary music, or when the performer is a computer program, which has not had the “appropriate” musical training. In these situations a more detailed representation, or set of instructions, is required (Selfridge-Field, 1997).

Each musical usage has a different set of requirements and thus a specific representation is needed for it. For example there are representations for controlling digital synthesizers and representations for analyzing the music structure of a piece of music with a computer. Given the currently developed computer applications we could group the used representations into three categories: (1) sound-related codes, (2) music notation codes, and (3) music data for analysis.

MIDI (Musical Instrument Data Interface) has been for some years, and still is, the most prevalent representation of music for computer applications. It was designed as a universal interconnection scheme for instrument controllers and synthesizers, thus in the category of sound-related codes, but through the years its usage has been extended to many other applications, even though it might not be the most appropriate one. Apart from MIDI, there are other representations at the sound level, such as Csound (Boullanger, 2000) that are also used for making music with software programs. But there is no definite solution to the music representation problem at the sound control level, a representation that should be able to express music at different levels of detail and at the same time useful for the control of all the available synthesis techniques.

Music notation codes are mainly used for computer representation and printing of traditional music, thus they are used by all the notation programs currently available. Looking at the quality obtained by these programs it is clear that the basic problems have mostly been solved. This is not the case for the codes used for music analysis, there is very little done in this area and it is not a simple problem to be solved. We need a representation from which structural analysis can be done, thus it has to show the relationships between the different music elements at different abstraction levels.

5. Performer

Despite being the performer a key element in the music making chain, there has been little formalization, or scientific studies, of what it really brings into that chain. Performance skills are learned in an intuitive way and thus they are quite elusive to scientific analysis. However in the past few years it has turned into a very active and fruitful field of research (Dannenberg, De Poli, 1998; Gabrielsson, 1999) with already some very promising results.

What makes a piece of music come alive is the performer's understanding of the structure and "meaning" of a piece of music, and his/her expression of this understanding via expressive performance. Hence the key research issue is to explain and quantify the principles that govern expressive performance. There are several strategies, complementary to each other, to approach such a problem:

- *Measuring performances.* Where we compare the parameters defined in a musical score with the ones obtained from a recording of the same score, either audio recording or some measured gesture data (normally MIDI). We can then try to formalize these differences with a set of performance rules or models. Such an approach requires analysis techniques from which to extract the musical parameters of the recordings and ways to compare them with the musical score data.
- *By an analysis-by-synthesis method.* Where we formalize the knowledge given by experimented performers using the perceptual feed-back of a sound synthesis system. Thus we can implement and test the validity of the formalization, which is generally a symbolic rules system. This has been the approach carried out by the KTH group in Stockholm (Bresin, 2000; Friberg, 1991).

- *Devising performance models with automatic methods.* The goal is to develop AI methods to come up with quantitative models of expressive performances. For example, machine learning algorithms are devised that search for systematic connections between structural aspects of the music and typical expression patterns (Widmer, 2001).

Current results on performance analysis give us some light into expressive aspects such as tempo and dynamics, but we are far from understanding most of what a performer really brings into the musical communication process.

6. Temporal Controls

One of the problems in studying performance issues is that it is very hard to differentiate the output of the performer from the output of the instrument being played. There is a strong coupling between the two, and in fact, in most cases, the two processes form a feed back loop impossible to understand one without the other. Generally, we can only record the output sound of the instrument, from which we will then have to differentiate the two types of data.

By placing sensors on the instrument it is possible to measure the action of the performer, such as the pressure of the fingers on a string. For example, in the case of the piano it is quite easy to measure the action of every single key of the keyboard and convert that to MIDI values. This is one of the reasons why most research on performance is being done on MIDI data extracted from piano performances. However, in self-sustained instruments, like bowed string or wind instruments, it is much harder to measure all the performance gestures and little work has been done on developing representations for this type of performance data.

In the context of electronic instruments it is feasible to separate the controlling aspect of a musical instrument from its sound producing capabilities. We can build controllers and interfaces to capture performance gestures and sound modules to produce sounds. With this division a staggering range of possibilities become available (Rowe, R. 1994; Winkler, 1998). Since the invention of the first electronic instruments there has been considerable research on developing new controllers with which to explore new creative possibilities (Chadabe, 1997) and communication protocols to interface the controllers with the sound generation devices (Selfridge-Field, 1997). Thus the concept of performance takes a new meaning and with it the concept of instrument. There is a huge open ground for research by considering the performer-instrument interface in the general framework of human-computer interaction.

7. Instrument

This is one of the best-defined elements of the chain and the one that yields the clearest scientific and technological problems. The understanding of the acoustics of the music instruments is a problem already introduced by the Greeks, with still fascinating challenges to be solved, for example, we still do not know why a Stradivarius sounds the way it does. In the Computer Music context we are also interested in inventing digital instruments based on computer models, or in modifying the sound of existing acoustical instruments by digital means, thus extending the creative possibilities of the composer, performer, and instrument builder.

The way to create digital instruments is by using synthesis techniques (Roads, 1996) and implementing them either in hardware or software. Traditionally, these techniques have been classified into: additive synthesis, subtractive synthesis, and non-linear synthesis. Additive synthesis is based on the sum of elementary sounds, each of which is generated by an oscillator. Subtractive synthesis is based on the complementary idea of filtering out parts of a complex

sound. The last group, non-linear synthesis, is a jumble in which a great number of techniques based on mathematical equations with non-linear behavior are included.

After many years trying to come up with “yet a new synthesis technique” the main research efforts in sound synthesis and instrument modeling are currently centered around two clear modeling problems. Either we want to model the sound source (physical modeling approach) or we want to model the perceived signal (spectral modeling approach).

With the physical modeling approach we generate sounds describing the behavior of the elements that make up a music instrument, such as strings, reeds, lips, tubes, membranes and resonant cavities. All these elements, mechanically stimulated, vibrate and produce disturbances, generally periodic, in the air that surrounds them. It is this disturbance that arrives to our hearing system and is perceived as sound. Historically, physical models have been carried out by means of very complex algorithms that can hardly work in real time with current technology. These implementations have been based on numerical integration of the equation that describes wave propagation in a fluid. Recently, more efficient solutions have been found for this problem and systems have begun to appear with interest for musicians (Smith, 1992).

Spectral models are based on the description of sound characteristics that the listener perceives. To obtain the sound of a string, instead of specifying the physical properties, we describe the timbre or spectral characteristics of the string sound. Then, sound generation is carried out from these perceptual data, thanks to diverse mathematical procedures developed in the last few decades. One advantage of these models is that techniques exist for analyzing sounds and obtaining the corresponding perceptual parameters. That is to say, by analyzing a specific sound we can extract its perceptual parameters. From the analysis, it is possible to synthesize the original sound again and the parameters can be modified in the process so that the resulting sound is new but maintains aspects of the analyzed sound (Serra, 1997).

8. Source Sound

The output of the instrument can be easily recorded with a microphone and processed with an analog to digital converter. Thus we get a direct representation of the sound as a one-dimensional pressure wave, represented digitally as a signal regularly sampled in time. There are many standard ways to store this information and also ways to visualize and analyze it. This is necessary in order to study the important physical attributes of the sound, such as its frequency, amplitude and timbre.

Another useful sound representation is obtained by decomposing the time waveform into its frequency components, thus obtaining a spectrum. Such a representation is based on the assumption that the waveform is stable during a certain length of time and it is generally expressed in polar coordinates as magnitude and phase values. To capture the time-varying characteristics of the sound we use a frame-based approach, thus the representation is a sequence of spectra. There are many different techniques for obtaining this representation from the time domain waveform, each one having a different set of compromises and being used in different applications. However the Fourier Transform is by far the most used approach.

There is a lot of research to come up with compressed representations of the signal, in order to reduce the memory needed for its storage or the bandwidth required for its transmission. But basically, the digital representation of the direct output of a music instrument is a solved problem.

9. Room

The space that the music is played into has as much impact on the listener's experience as the instruments used. But, even though the science of room acoustics is several centuries old, nearly all the concert halls still have acoustic problems. Most of the individual issues involved in room acoustics are quite well known, nevertheless the complexity of real spaces is so huge that we are still far from being able to design the "perfect" concert hall.

In the context of Computer Music we are especially interested in simulating spaces with computer models and being able to control the location of our sound sources, or instruments, inside these spaces (Begault, 1994). We can characterize the reverberation of a space by its impulse response, or from a parameterization of that signal. From this characterization we can then create digital reverberators using different signal processing strategies based on digital filters and delay lines. But given the complexity of real spaces, the creation of natural reverberations is still a great challenge. At the same time the simulation of localization and movement of sound through space brings many interesting problems. Depending on the sound reproduction situation the simulation of localization and movement cues is done in different ways. Whether we have a standard stereo system, a multichannel system or headphones the strategy will be completely different. Much of the work on sound spatialization is based on psychoacoustics' studies, thus reproducing the space perceptual cues.

10. Sound Field

Many of the recent improvements made to the commercial sound systems have been incorporated in order to have the possibility to preserve into a sound recording a given sound field and to be able to reproduce it (Begault, 1994). Thus, giving the listener the sensation of being in the "original" acoustic space, a 3D sound environment. From the old monophonic recordings to the current surround systems there has been an enormous evolution. Some of the landmarks have been: Two-channel recordings, Three-channel systems, Four-channel surround systems, Binaural systems, Ambisonics (Malham and Myatt, 1995), and Motion Picture Sound (ex. Dolby Stereo).

11. Listener

From a medical point of view, the physiology of the human hearing is quite well understood, but the relationship between acoustics, music structure, cognition, and emotion is an ongoing study far from being understood. There are a lot of issues involved in the process that goes from the sound waveform entering our ear to the music sensation we get from that sound. We understand some of the low level (hearing system) issues but we are far from understanding the cognition issues. There are many areas involved, related and interrelated to the different qualities and aspects of auditory perception, some of which are: basic auditory processes, low-high grouping mechanisms, timbre, pitch, time and rhythm perception (Deutsch, 1998).

In auditory modeling the aim is to find mathematical models that represent some physiological behavior or some perceptual aspects of human hearing (Hawkins, McMullen, Popper, Fay, 1996). Thus with a good model we can analyze audio signals in a similar way that our brain does.

The computational models of the ear (Lyon, 1982) generally give particular attention the behavior of the cochlea, the most important part of the inner ear, that acts substantially as a non-linear filter bank. But one of the better known issues is the filtering effect of the head and the outer ear, which is critical for the spatial perception of sound (Blauert, 1983). This processing effect is generally modeled by the Head Related Transfer Functions, HRTF, which are represented by pole/zero models, series expansions or structural models.

Beyond the low level physiological aspects there are many research topics related to disciplines such as psychoacoustics, psychology or cognitive musicology. For example, to deal with the sound environment the auditory system has to extract the sound components we need from the sound signal reaching the ear and to construct an auditory world. This function of the auditory system is known as auditory scene analysis (Bregman, 1990). It is still a quite unknown process but considerable work is focused on its simulation by computer.

Another active field is related to our ability to distinguish and categorize sounds. The classification of musical sound into different timbre groups has received much attention in the perceptual literature in the past, but the principal components of timbre still remain elusive. Existing theories of timbre are derived from perceptual experiments, such as the well known multi-dimensional scaling experiments of Grey (Grey, 1976) in which the cognitive relationships amongst a group of sounds are represented geometrically. Current alternative approaches are based on machine learning algorithms or statistical analysis (Herrera, Amatriain, Batlle, Serra, 2000).

At the musical cognition level there has been several proposed theories such as Narmour's implication/realization model (Narmour, 1990). It proposes a theory of cognition of melodies based on simple structures that characterize patterns of melodic implications that constitute the basic units of the listener perception. Lerdahl and Jackendoff's proposed a generative theory of tonal music (Lerdahl and Jackendoff, 1993) which offers an alternative approach to understanding melodies based on a hierarchical structure of musical cognition.

12. Perception and Cognition

A perceptually meaningful sound representation should be based on log-frequency in order to reflect the relative salience of low frequency components with respect to the high frequencies. There must be also temporal information that gives precedence to the transient portion of the signal. The standard Fourier representations do not fulfill these requirements and thus there has been considerable research into alternative representations. We could just mention a few:

- The correlogram (Slaney and Lyon, 1993) allows us to see where energy is located in a log frequency, but also the value of the autocorrelation lag for which the signals of the cochlear channels have the same periodicity.
- Mel-warping of spectra is commonly used in cepstral-based speech processing and since the Mel scale was derived from human psychophysics, the resulting frequency scale is cochlear like.
- The cochleogram is the direct output of a cochlear filterbank model.
- The field of Computational Auditory Scene Analysis is emerging with new, non FFT-based, representations of audio with a view to solving difficult auditory scene analysis problems. Frequency components are grouped using gestalt principles such as synchrony of onset and temporal proximity as well as psycho-acoustic principles such as harmonicity and critical band masking effects.

13. Conclusions

The complexity of the music communication chain is too huge to be approached by any single discipline. Throughout this short article we have referenced contributions from such disciplines as: Music, Computer Science, Electrical Engineering, Psychology, Physics. This is precisely one

of the virtues of the Computer Music field and what Moore (Moore, 1990) tries to express in the diagram shown in Figure 2.

The number of open problems is still vast and we have only pointed out some of them. However it is clear that the interdisciplinary approach expressed in this article is the only way to tackle most of them.

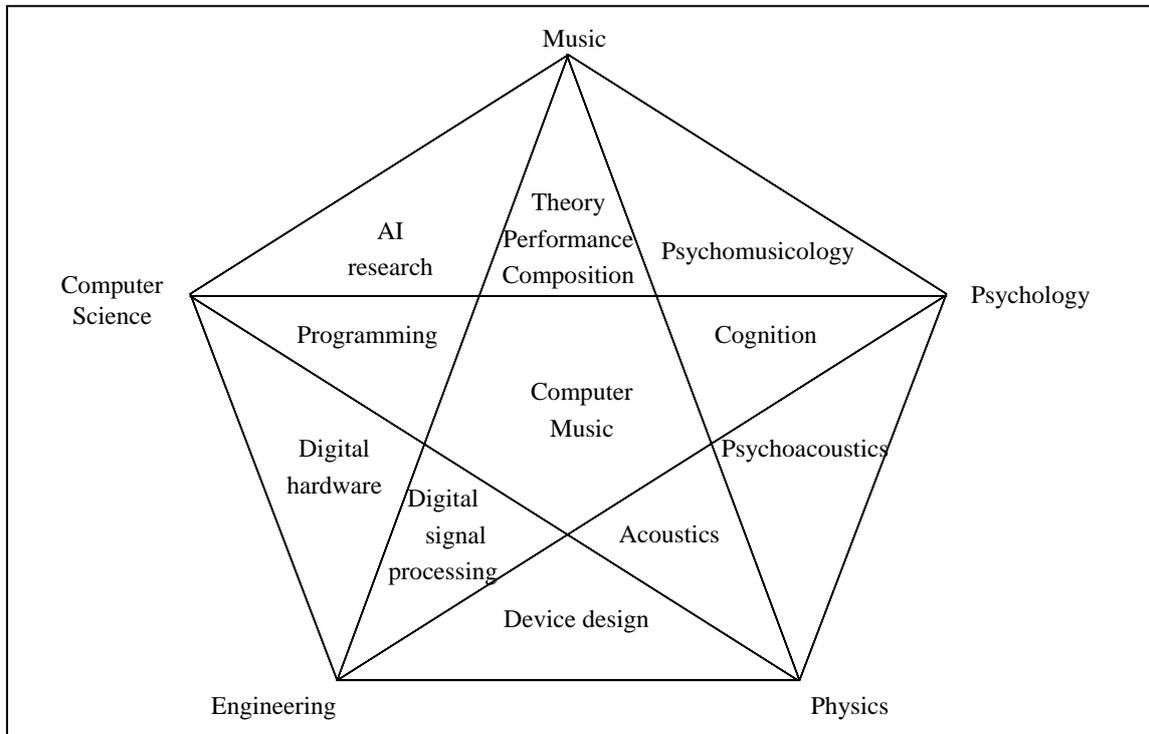


Figure 2: Diagram of the interdisciplinarity of the Computer Music field as proposed by Moore.

References

- Begault, D. R. 1994. *3-D Sound For Virtual Reality and Multimedia*. Academic Press.
- Boulanger, R. C. 2000. *The Csound Book: Perspectives in Software Synthesis, Sound Design, Signal Processing, and Programming*, Cambridge, MA: MIT Press.
- Blauert, J. 1983. *Spatial Hearing*. Cambridge, MA: MIT Press.
- Bregman, A.S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge, MA: MIT Press.
- Bresin, R. 2000. *Virtual Virtuosity: Studies in Automatic Music Performance*, Ph.D. Thesis. KTH, Stockholm, Sweden.
- Brown, Andrew R. 1998. "Your Friend, the Algorithm". *Music Forum*. Vol.4:6. Sydney: Music Council of Australia. pp. 26-29.
- Chadabe, J. 1997. *Electric Sound: The Past and Promise of Electronic Music*, Prentice Hall.
- Cook, P. R. 1999. *Music Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. Cambridge, MA: MIT Press.
- Cope, D. 1991. *Computers and Musical Style* (The Computer Music and Digital Audio Series, Vol. 6). Madison, Wisconsin: A-R Editions.
- Dannenberg, R.; G. De Poli. 1998. "Synthesis of Performance Nuance", special issue of *Journal of New Music Research*, Vol. 27, number 3, September 1998.

- Deutsch, D. 1998. *The Psychology of Music*, Academic Press Series in Cognition and Perception, 2nd. edition.
- Dodge, C.; T. A. Jerse. 1996. *Computer Music*, Schirmer Books, New York. (Second edition)
- Fletcher, N. H.; T. D. Rossing. 1991. *The Physics of Musical Instruments*, New York: Springer-Verlag.
- Friberg, A. 1991. "Generative Rules for Music Performance: A Formal Description of a Rule System", *Computer Music Journal*, 15, 56-71.
- Gabrielsson, A. 1999. "The Performance of Music", in D. Deutsch (Ed.) *The Psychology of Music*. San Diego: Academic Press [second edition], 501-602.
- Grey, J. M. 1976. "Multidimensional perceptual scaling of musical timbres", *Journal of the Acoustical Society of America*, 61(5):1270-1277.
- Hawkins, H. L.; T. A. McMullen; A. N. Popper; R. R. Fay. 1996. *Auditory Computation*, Springer-Verlag.
- Herrera, P; X. Amatriain; E. Batlle; X. Serra. 2000. "Towards Instrument Segmentation for Music Content Description: a Critical Review of Instrument Classification Techniques" *Proceedings of International Symposium on Music Information Retrieval*.
- Hiller, L.; L. Isaacson. 1959. *Experimental Music*, New York: McGraw-Hill Book Company, Inc.
- Lerdahl, F.; R. Jackendoff. 1993. "An overview of hierarchical structure in music", S. M. Scwanaver and D. A. Levitt, editors, *Machine Models of Music*. Reproduced from Music Perception.
- Narmour, E. 1990. *The analysis and cognition of basic melodic structures: the implication-realization model*, University of Chicago Press.
- Loy, G. 1989. "Composing with Computers: A survey of Some Compositional Formalisms and Music Programming Languages", *Current Directions in Computer Music Research*, edited by. M. V. Mathews and J. R. Pierce. MIT Press, 1989.
- Lyon, R. F. 1982. "A Computational Model of Filtering, Detection, and Compression in the Cochlea", *Proceedings of IEEE-ICASSP-82*, pp. 1282-1285.
- Malham, D.G.; A. Myatt. 1995. "3-D Sound Spatialization using Ambisonic Techniques", *Computer Music Journal*, 19;4, pp 58-70.
- McAdams, S. 1996. "Audition: Cognitive Psychology of Music", R. Llinas & P. Churchland (eds.), *The Mind-Brain Continuum*, pp. 251-279, Cambridge, MA: MIT Press.
- Moore, F. R. 1990. *Elements of Computer Music*, New Jersey: Prentice-Hall, Inc.
- Roads, C. 1996. *The Computer Music Tutorial*, Cambridge, MA: MIT Press.
- Rowe, R. 1994. *Interactive Music Systems - Machine Listening and Composing*, Cambridge, MA: MIT Press.
- Schillinger, J. 1948. *The Mathematical Basis of the Arts*, New York: The Philosophical Library.
- Selfridge-Field, E. 1997. *Beyond MIDI: The Handbook of Musical Codes*, Cambridge, MA: MIT Press.
- Serra, X. 1997. "Musical Sound Modeling with Sinusoids plus Noise", in G. D. Poli, A. Picialli, S. T. Pope, and C. Roads, editors, *Musical Signal Processing*, Swets & Zeitlinger Publishers.
- Slaney, M.; R.F. Lyon. 1993. "On the Importance of Time: A Temporal Representation of Sound", in Cooke M., Beet S. and Crawford M. (Eds.): *Visual Representations of Speech Signals*, Wiley & Sons, Chichester, pp. 95-115.
- Smith J. O. 1992. "Physical modeling using digital waveguides", *Computer Music Journal*, 16:74-87.
- Widmer, G. 2001. "Machine Learning and Expressive Music Performance", *AI Communications*.
- Winkler, T. 1998. *Composing interactive music: techniques and ideas using MAX*, Cambridge MA: MIT Press.

- Xenakis, I. 1971. *Formalized Music*, Bloomington: Indiana University Press.