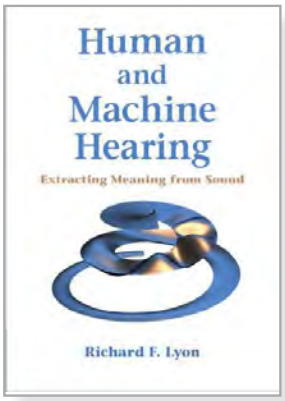
Human and Machine Learning: Extracting Meaning From Sound



Author: Richard F. Lyon
Publisher: Cambridge University Press, Cambridge, UK, 2017, 567 pp.
Price: £54.99
ISBN: 978-1-107-007536

Richard (Dick) Lyon, a Principal Research Scientist at Google, is well known for his work on models of the auditory system, particularly cochlear models, and for developing analog and digital implementations of those models, in hardware and software. Although he has published extensively, this is, as far as I know, his first book. Although the book can be purchased in hardback form, all parts of the book are available free from the web site of the publisher—acousticstoday.org/lyonbook.

This is a very long and comprehensive book, covering many aspects of hearing, including the analysis of sounds in the peripheral auditory system, human auditory perception, methods of analyzing sounds, linear systems, digital representations of signals, digital signal processing, nonlinear systems, automatic gain control, models of the cochlea, models of neural transduction, the representation of sound in the auditory nerve and cochlear nucleus, binaural hearing, auditory scene analysis, and several aspects of machine learning and its uses for sound classification, speech recognition, searching for sounds, and melody matching. This is a very impressive range of topics for one person to cover.

The general approach of the book is unusual, perhaps even a touch eccentric. Most chapters contain a mix of historical material (including historical quotations) and didactic material. While the historical material is interesting and entertaining, those who wish to learn about a specific topic may find it frustrating to have to read through this material in order to get to the real "meat." In the Preface, Lyon states "… interest in sound comes from people of many different disciplines, with complementary backgrounds and sometimes incompatible terminology and concepts. I want all of these people as my audience." He adds that the target audience includes "electrical engineers, computer scientists, physicists, physiologists, audiologists, musicians, psychologists, and others…." This is an ambitious goal and I think that Lyon achieves it only partially. Some parts of the book will be very heavy going for audiologists and musicians without a technical background, and students or researchers who want to learn about a specific topic may find it frustrating that the relevant information is spread across chapters and mixed with historical detail. Nevertheless, the book does contain a wealth of interesting information, and there is something of interest for the great majority of people who are interested in hearing.

Generally, the text is authoritative and comprehensive. However, there are some parts that are a bit imprecise. For example, Lyon is not always careful to distinguish between frequency and fundamental frequency (physical attributes of sound) on the one hand and pitch (a subjective attribute of sound) on the other hand. On p. 25 he talks about "the coding of perceived pitch frequency" and says that "the pitch of a sinusoid is equal to its frequency," whereas it would be more accurate to say that the pitch is primarily determined by the frequency, although sound level also has a small influence. The phon (the unit of loudness level) is initially not precisely defined, but is described as "… just a dB scale of intensity that is warped to correlate better with loudness for frequencies away from 1 kHz." In the section on "Hearing Aids and Cochlear Implants" Lyon states that modern hearing aids "make decisions about what is the signal of interest and present it in a way that makes up for the listener's particular hearing deficit." While this is a dream of many researchers, hearing aids are not yet capable of deciding what signals are of interest and they are far less than perfect in compensating for specific hearing deficits. Also, while the coverage is generally comprehensive, I found it surprising that the book does not mention the work of Houtgast (1972, 1973) on lateral suppression in hearing or the work of Dau and colleagues on the concept of the modulation filter bank (Dau et al., 1997a,b).

Despite these minor quibbles this book is an impressive achievement, presenting a huge range of material on human and machine hearing in a comprehensive manner. I would not recommend it as a text book, but it should be a valuable resource for any researcher in the field of hearing.

References

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997a). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrowband carriers. *The Journal of the Acoustical Society of America* 102, 2892–2905. https://doi.org/10.1121/1.420344.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997b). Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *The Journal of the Acoustical Society of America* 102, 2906–2919. https://doi.org/10.1121/1.420345.

Houtgast, T. (1972). Psychophysical evidence for lateral inhibition in hearing. T*he Journal of the Acoustical Society of America* 51, 1885–1894. https://doi.org/10.1121/1.1913048.

Houtgast, T. (1973). Psychophysical experiments on "tuning curves" and "two-tone inhibition," *Acustica* 29, 168–179.

Review by:
Brian C. J. Moore
Department of Psychology, University of Cambridge
Downing Street, Cambridge CB2 3EB, UK

[Published online August 2, 2018, The Journal of the Acoustical Society of America, *144(2)]*